Paidós
Básica

13

Carl G. Hempel

¡eureka!

La explicación científica

Estudios sobre la filosofía de la ciencia



La explicación científica

Paidós Básica

Últimos títulos publicados:

33. G. Duby - *Europa en la Edad Media*
34. C. Lévi-Strauss - *La alfarera celosa*
35. J. W. Vander Zanden - *Manual de psicología social*
36. J. Piaget y otros - *Construcción y validación de las teorías científicas*
37. S. J. Taylor y R. Bogdan - *Introducción a los métodos cualitativos de investigación*
38. H. M. Feinstein - *La formación de William James*
39. H. Gardner - *Arte, mente y cerebro*
40. W. H. Newton-Smith - *La racionalidad de la ciencia*
41. C. Lévi-Strauss - *Antropología estructural*
42. L. Festinger y D. Katz - *Los métodos de investigación en las ciencias sociales*
43. R. Arrillaga Torrens - *La naturaleza del conocer*
44. M. Mead - *Experiencias personales y científicas de una antropóloga*
45. C. Lévi-Strauss - *Tristes trópicos*
46. G. Deleuze - *Lógica del sentido*
47. R. Wuthnow - *Análisis cultural*
48. G. Deleuze - *El pliegue. Leibniz y el barroco*
49. R. Rorty, J. B. Schneewind y Q. Skinner - *La filosofía en la historia*
50. J. Le Goff - *Pensar la historia*
51. J. Le Goff - *El orden de la memoria*
52. S. Toulmin y J. Goodfield - *El descubrimiento del tiempo*
53. P. Bourdieu - *La ontología política de Martin Heidegger*
54. R. Rorty - *Contingencia, ironía y solidaridad*
55. M. Cruz - *Filosofía de la historia*
56. M. Blanchot - *El espacio literario*
57. T. Todorov - *Crítica de la crítica*
58. H. White - *El contenido de la forma*
59. F. Rella - *El silencio y las palabras*
60. T. Todorov - *Las morales de la historia*
61. R. Koselleck - *Futuro pasado*
62. A. Gehlen - *Antropología filosófica*
64. R. Rorty - *Ensayos sobre Heidegger y otros pensadores contemporáneos*
65. D. D. Gilmore - *Hacerse hombre*
66. C. Geertz - *Conocimiento local*
67. A. Schütz - *La construcción significativa del mundo social*
68. G. E. Lenski - *Poder y privilegio*
69. M. Hammersley y P. Atkinson - *Etnografía. Métodos de investigación*
70. C. Solís - *Razones e intereses*
71. H. T. Engelhardt - *Los fundamentos de la bioética*
72. E. Rabossi y otros - *Filosofía de la mente y ciencia cognitiva*
73. J. Derrida - *Dar (el) tiempo. I. La moneda falsa*
74. R. Nozick - *La naturaleza de la racionalidad*
75. B. Morris - *Introducción al estudio antropológico de la religión*
76. D. Dennet - *La conciencia explicada. Una teoría interdisciplinar*
79. R. R. Aramayo, J. Muguerza y A. Valdecantos - *El individuo y la historia*

Carl G. Hempel

La explicación científica

Estudios sobre la filosofía de la ciencia



Título original: *Aspects of scientific explanation and other essays in the philosophy of science*
Publicado en inglés por The Free Press, Nueva York

Traducción: M. Frassineti de Gallo (cap. 8)
Néstor Míguez (caps. 1-6 y 12)
Irma Ruiz Aused (caps. 7 y 11)

Cubierta de Mario Eskenazi

1.ª edición, 1979
2.ª reimpresión, 1996

ISBN: 84-7509-497-X
Depósito legal: B-1.996/1996

Impreso en Novagràfik, S.L.,
Puigcerdà, 127 - 08019 Barcelona

Impreso en España - Printed in Spain

Indice

A Diane

Prefacio

Los ensayos reunidos en este volumen se refieren a cuatro temas principales de la filosofía de la ciencia, por lo cual han sido agrupados bajo los rótulos "Confirmación, inducción y creencia racional", "Concepciones de significación cognoscitiva", "Estructura y función de los conceptos y teorías científicos" y "La explicación científica".

Todos los capítulos, excepto uno, son versiones revisadas de artículos ya aparecidos, según se indica en las notas de pie de página referentes a su origen. El más largo de los ensayos, del cual toma su título esta compilación, fue escrito expresamente para este volumen. Ofrece un estudio independiente de la explicación científica, en el que se incluye un nuevo examen del concepto de explicación por leyes abarcantes, parcialmente desarrollado en dos ensayos anteriores, publicados aquí como capítulos 9 y 10. El ensayo del título también trata con algún detalle de la explicación por leyes estadísticas, tema que sólo había recibido una breve consideración en los artículos anteriores. El análisis de la explicación estadística que aquí presentamos difiere en aspectos importantes de un estudio anterior sobre la cuestión, publicado en 1962 y registrado en la bibliografía pero no incluido en este volumen.

Aunque los artículos 9 y 10 se superponen en parte con el ensayo del título, los reproducimos aquí porque se los ha discutido mucho en la bibliografía reciente sobre la explicación, de modo que valía la pena ponerlos a disposición del público para facilitar su consulta y también porque buena parte del material de esos artículos no está incluido en el ensayo del título.

Si bien considero aún que las ideas centrales de los ensayos reproducidos son básicamente correctas, he cambiado de opinión, como es natural, en lo referente a varios puntos de detalle. Cuando lo juzgué apropiado, señalé tales cambios en notas al pie con la indicación "Agregado en 1964" o en los Apéndices que añadí a tres de los artículos. También efectué, sin indicarlos, cambios estilísticos, supresión de pasajes que no agregaban nada a la argumentación y corrección de errores de menor importancia.

En los Apéndices mencionados comenté también algunos desarrollos recientes en el análisis filosófico de los problemas centrales y agregué

algunas reflexiones posteriores. Pero no intenté actualizar las bibliografías de todos los artículos reproducidos, ya que habría carecido de objeto registrar publicaciones recientes sin examinar su contenido.

Como traté de ponerlo en claro en los lugares apropiados de estos ensayos y en las notas agregadas y los Apéndices, me he beneficiado mucho con la labor de otros, con discusiones y críticas de mis escritos que han aparecido impresas y con el intercambio personal de ideas con amigos, colegas y estudiantes: a todos estos benefactores intelectuales míos les estoy agradecido.

Varios de estos ensayos fueron escritos durante los meses de verano en el retiro, con aire acondicionado, de la casa de mis viejos amigos Paul y Gabrielle Oppenheim, en Princeton. A Paul Oppenheim, con quien he discutido cuestiones filosóficas durante muchos años, le agradezco también el permitirme reproducir aquí uno de los artículos que escribimos juntos. Trabajé en algunos de los otros ensayos durante un año, de 1959 a 1960, como Fulbright Research Fellow, en Oxford. Finalmente, una licencia sabática de la Universidad de Princeton, junto con una beca para 1963-1964 en este refugio erudito que es el Centro de Estudios Avanzados de Ciencias de la Conducta, me permitió escribir el ensayo que dio título al libro y revisar los artículos anteriores para su publicación.

Agradezco mucho a los editores que me han permitido reproducir los artículos y extractos incluidos en este volumen.

Dedico este libro, con agradecimiento, a mi esposa. Su afectuoso estímulo e inquebrantable apoyo habrían merecido una ofrenda mejor.

C. G. H.

Stanford, California

Primera Parte

CONFIRMACION, INDUCCION Y CREENCIA RACIONAL

CAPÍTULO I

Estudios sobre la lógica de la confirmación [1]

1. *Objetivos de este estudio* [2]

La característica definitoria de un enunciado empírico es su capacidad de ser sometido a prueba mediante la confrontación con hallazgos experimentales, es decir, con los resultados de experimentos apropiados o de la observación dirigida. Esta característica distingue los enunciados que tienen contenido empírico tanto de los enunciados de las ciencias formales, la lógica y la matemática, que no necesitan de ningún test experiencial para su validación, como de las formulaciones de la metafísica transempírica, que no admite ningún test semejante.

La testabilidad a la que aquí nos referimos debe ser entendida en el sentido amplio de "testabilidad en principio" o "testabilidad teórica"; muchos enunciados empíricos no pueden ser testados por ahora, por razones prácticas. Decir que un enunciado de este tipo es testable en principio significa que es posible indicar exactamente qué hallazgos experimentales, si se los obtuviera realmente, constituirían elementos de juicio favorables a él, y qué hallazgos o "datos", como diremos para mayor brevedad, constituirían elementos de juicio desfavorables. En otras palabras, decimos que un enunciado es testable en principio si es posible describir el tipo de datos que lo confirmarían o desconfirmarían.

Los conceptos de confirmación y desconfirmación, tal como aquí los entendemos, son claramente más amplios que los de verificación y refuta-

[1] Este capítulo es una reimpresión, con algunos cambios, realizada con la amable autorización del director de *Mind*, donde apareció en el volumen 54, págs. 1-26 y 97-121 (1945).

[2] El siguiente análisis de la confirmación fue sugerido y estimulado en gran medida por un estudio en común de ciertos problemas más generales que fueron planteados por el doctor Paul Oppenheim y que he investigado con él durante varios años. Estos problemas se refieren a la forma y función de las leyes científicas y a la metodología comparada de las diferentes ramas de la ciencia empírica.

En mi estudio de los aspectos lógicos de la confirmación, es grande el provecho que he obtenido de discusiones con los profesores R. Carnap, A. Tarski y, en particular, con el doctor Nelson Goodman, a quien debo varias sugerencias valiosas que indicaré más adelante.

Una exposición detallada de los aspectos más técnicos del análisis de la confirmación presentado en este ensayo se encontrará en mi artículo "A Purely Syntactical Definition of Confirmation", *The Journal of Symbolic Logic*, vol. 8 (1943).

ción concluyentes. Por ejemplo, ninguna cantidad finita de elementos de juicio experimentales puede verificar de modo concluyente una hipótesis que exprese una ley general, como la ley de gravitación, que abarca una infinidad de casos potenciales, muchos de los cuales pertenecen al futuro aún inaccesible o al pasado ya irrecuperable; pero un conjunto finito de datos atinentes a la cuestión puede estar "de acuerdo con" la hipótesis y, así, constituir elementos de juicio confirmatorios de ella. Análogamente, una hipótesis existencial concerniente, digamos, a la existencia de un elemento químico aún desconocido con ciertas características específicas, no puede ser refutada de manera concluyente por una cantidad finita de elementos de juicio que no "den sustento" a la hipótesis; pero, en ciertas condiciones, tales datos desfavorables pueden ser considerados como debilitantes de la hipótesis en cuestión, o como elementos de juicio desconfirmatorios de ella.[3]

Si bien en la investigación científica concreta los juicios relativos al carácter confirmatorio o desconfirmatorio de los datos de la experiencia obtenidos en el test de una hipótesis a menudo son expresados sin vacilación y con un amplio acuerdo de opinión, no puede decirse que tales juicios se basen en una teoría explícita que suministre criterios generales de confirmación o desconfirmación. A este respecto, la situación es comparable con la forma en que se realizan las inferencias deductivas en la práctica de la investigación científica: a menudo se las realiza también sin referencia a un sistema explícitamente formulado de reglas de inferencia lógica. Pero mientras que la lógica formal puede suministrar y ha suministrado criterios válidos de deducción, en cambio no parece haber hasta ahora ninguna teoría satisfactoria que brinde criterios generales de confirmación o desconfirmación.

En este ensayo trataremos de colocar los cimientos de una teoría de este tipo. Después de una breve revista de la significación y el estado actual del problema, me propongo presentar un análisis crítico detallado de algunas concepciones comunes sobre la confirmación y la desconfirmación, para luego construir definiciones explícitas de estos conceptos y formular algunos principios básicos de lo que podría llamarse la lógica de la confirmación.

2. *Significación y estado actual del problema*

La elaboración de una teoría general de la confirmación bien puede considerarse como uno de los más urgentes objetivos de la metodología actual de las ciencias empíricas. En verdad, pareciera que un análisis preciso del concepto de confirmación es una condición necesaria para hallar soluciones adecuadas de varios problemas fundamentales concernientes a la estructura lógica del proceder científico. Hagamos una breve revista de los más importantes de estos problemas.

[3] Este punto, así como el de la posibilidad de la verificación y la refutación concluyentes, será examinado con mayor detalle en la sección 10 de este capítulo.

a) En el examen del método científico, el concepto de elementos de juicio atinentes a una cuestión desempeña un papel importante. Y si bien ciertas doctrinas inductivistas del proceder científico parecen suponer que los elementos de juicio atinentes a una cuestión, o los datos atinentes al caso, pueden ser reunidos en el contexto de una indagación anterior a la formulación de cualquier hipótesis, no cabe duda de que, a poco que se reflexione sobre ello, el de atingencia es un concepto relativo; puede decirse que los datos experienciales son o no atinentes a una hipótesis dada, y es ésta la que determina el tipo de datos o elementos de juicio que le son atinentes. En verdad, un hallazgo empírico es atinente a una hipótesis si y sólo si constituye un elemento de juicio favorable o desfavorable a ella; en otras palabras, si confirma o desconfirma la hipótesis. Así, una definición precisa de atingencia presupone un análisis de la confirmación y de la desconfirmación.

b) Un concepto íntimamente relacionado con el anterior es el de caso (*instance*) de una hipótesis. Habitualmente se presenta el llamado método de inferencia inductiva como el paso de casos específicos a una hipótesis general, de la cual cada uno de los fenómenos específicos es un "caso" en el sentido de que está de acuerdo con la hipótesis general en cuestión y, por ende, constituye un elemento de juicio confirmatorio de ella.

De este modo, todo examen de la inducción que se refiera al establecimiento de hipótesis generales sobre la base de casos particulares presenta todas las dificultades lógicas —pronto las expondremos— que rodean el concepto de confirmación. Un análisis preciso de este concepto, pues, es una condición necesaria para una enunciación clara de las cuestiones implicadas en el complejo problema de la inducción y de las ideas sugeridas para su solución, sean cuales fueren sus méritos o deméritos teóricos.

c) Otra cuestión habitualmente vinculada con el estudio del método científico es la búsqueda de "reglas de inducción". Hablando en términos generales, tales reglas nos permitirían inferir, a partir de un determinado conjunto de datos, la hipótesis o generalización que mejor se adecua a todos los datos particulares del conjunto dado. Pero esta manera de plantear el problema supone una concepción errónea; pues, si bien el proceso de invención por el cual se realizan los descubrimientos científicos está, por lo general, *guiado y estimulado psicológicamente* por el conocimiento anterior de hechos específicos, sus resultados no están *determinados lógicamente* por ellos; los medios por los que se descubren hipótesis o teorías científicas no pueden ser formulados en un conjunto de reglas generales de la inferencia inductiva.[4] Una de las consideraciones fundamentales que llevan a esta conclusión es la siguiente: Tómese una teoría científica como la teoría atómica de la materia. Es posible descri-

[4] Véase la lúcida presentación de este punto en *Logik der Forschung* (Viena, 1935) de Karl Popper, en particular las secciones 1, 2, 3 y 25, 26 y 27; véanse también las observaciones de Albert Einstein en su conferencia *On the Method of Theoretical Physics* (Oxford, 1933), 11 y 12. En este contexto también presenta interés el examen crítico de la inducción realizado por H. Feigl en "The Logical Character of the Principle of Induction", *Philosophy of Science*, vol. 1 (1934).

bir los elementos de juicio sobre los cuales reposa en términos que se refieran a fenómenos directamente observables, a saber, a ciertos aspectos macroscópicos de los diversos datos experimentales y observaciones atinentes a la teoría. Por otra parte, la teoría misma contiene un gran número de términos no-observacionales sumamente abstractos, tales como "átomo", "electrón", "núcleo", "disociación", "valencia" y otros, ninguno de los cuales figura en la descripción de los datos observacionales. Una regla de inducción apropiada, pues, debería suministrar, para éste y otros casos concebibles, criterios mecánicamente aplicables que permitan determinar sin ambigüedad, y sin basarse para nada en la inventiva o en el conocimiento científico adicional de quien los usa, todos esos nuevos conceptos abstractos que es necesario crear para la formulación de la teoría destinada a explicar los elementos de juicio dados. Es claro que este requisito no puede ser satisfecho por ningún conjunto de reglas, por ingeniosamente que se las conciba. No puede haber en absoluto reglas de inducción en el sentido anterior, y la exigencia de tales reglas reposa en una confusión de los problemas lógicos con los psicológicos. Aquello que determina la corrección de una hipótesis no es la manera como se llegue a ella (hasta puede haberla sugerido un sueño o una alucinación), sino su resistencia a las pruebas, es decir, a la confrontación con datos observacionales atinentes al caso. Por consiguiente, la búsqueda de reglas de inducción en el sentido original de cánones del descubrimiento científico debe ser reemplazada, en la lógica de la ciencia, por la búsqueda de criterios generales objetivos que determinen: (A) si puede decirse de una hipótesis *H* que está corroborada por un conjunto dado de elementos de juicio *E*, y —si es posible— (B) en qué grado. Este enfoque difiere esencialmente de la concepción inductivista del problema en que éste presupone no solamente *E*, sino también *H*, como dados, y luego trata de determinar una cierta relación lógica entre ellos. Es posible relacionar las dos partes de este último problema en términos un poco más precisos, del siguiente modo:

(A) Dar definiciones precisas de los dos conceptos relacionados no-cuantitativos de confirmación y desconfirmación; es decir, definir el significado de las frases "*E* confirma *H*" y "*E* desconfirma *H*". (Cuando *E* no confirma ni desconfirma *H*, diremos que *E* es neutral, o carente de atingencia, con respecto a *H*.)

(B) (1) Establecer criterios que definan un concepto métrico, el de "grado de confirmación de *H* con respecto a *E*", cuyos valores sean números reales, o en caso de que esto no sea posible:

(2) Establecer criterios que definan dos conceptos relacionales, los de "más confirmado que" e "igualmente confirmado que", los cuales permitan una comparación no métrica de hipótesis (cada una de las cuales tenga asignado un conjunto de elementos de juicio) con respecto a la extensión de su confirmación.

Es interesante el hecho de que el problema B haya recibido mucha más atención en la investigación metodológica que el problema A; en particular, puede considerarse que las diversas teorías acerca de la llamada probabilidad de las hipótesis conciernen a este problema comple-

jo; aquí hemos adoptado[5] la expresión más neutra de "grado de confirmación", en lugar de "probabilidad", porque esta última se usa en la ciencia con un sentido técnico definido que supone la referencia a la frecuencia relativa de la aparición de un suceso dado en una sucesión, y por lo menos aún está en discusión si el grado de confirmación de una hipótesis puede definirse, en general, como una probabilidad en este sentido estadístico.

Las teorías que tratan de la probabilidad de las hipótesis pueden clasificarse en dos grupos principales: las teorías "lógicas", que conciben la probabilidad como una relación lógica entre oraciones (o proposiciones, pues no siempre está claro a qué se alude)[6] y las teorías "estadísticas", que interpretan, en esencia, la probabilidad de una hipótesis como el límite de la frecuencia relativa de sus casos confirmatorios, entre todos los casos atinentes a la cuestión.[7] Ahora bien, es un hecho notable que ninguna de las teorías del primer tipo elaboradas hasta ahora suministra una definición explícita general de la probabilidad (o grado de confirmación) de una hipótesis *H* con respecto a un conjunto de elementos de juicio *E*; todas ellas se limitan esencialmente a la construcción de un sistema postulacional no interpretado de probabilidad lógica.[8] Por esta razón, estas teorías no ofrecen una solución completa del problema B. El enfoque estadístico, en cambio, si tiene éxito, suministraría una definición numérica explícita del grado de confirmación de una hipótesis; esta definición estaría formulada en términos del número de casos confirmatorios y desconfirmatorios de *H* que constituyen el conjunto de elementos de juicio *E*. Así, una condición necesaria de una interpretación adecuada de los grados de confirmación como probabilidades estadísticas es el establecimiento de criterios precisos de confirmación y desconfirmación; en otras palabras, la solución del problema A.

Sin embargo, a pesar de su gran ingeniosidad y sugestividad, los

[5] Siguiendo el uso de R. Carnap en "Testability and Meaning", *Philosophy of Science*, vols. 3 (1936) y 4 (1937), en particular el capítulo 8.

[6] Este grupo incluye la labor de autores tales como Janina Hosiasson-Lindenbaum [véase, por ejemplo, su artículo "Induction et analogie: Comparaison de leur fondement", *Mind*, vol. 50 (1941)], H. Jeffreys, J. M. Keynes, B. O. Koopman, J. Nicod, St. Mazurkiewicz y F. Waismann. Se hallará un breve examen de este concepto de la probabilidad en Ernest Nagel, *Principles of the Theory of Probability* (International Encyclopedia of Unified Science, vol. I, nº 6, Chicago, 1939), en particular secciones 6 y 8.

[7] El principal defensor de esta concepción es Hans Reichenbach; véase en particular "Ueber Induktion und Wahrscheinlichkeit", *Erkenntnis*, vol. 5 (1935), y *Experience and Prediction* (Chicago, 1938), cap. V.

[8] (Agregada en 1964.) Desde que fue escrito este capítulo R. Carnap ha desarrollado una teoría de la lógica inductiva que permite explícitamente, para lenguajes formalizados de cierto tipo, definir —sin usar la noción cualitativa de caso confirmatorio— un concepto cuantitativo de grado de confirmación que tiene las características formales de una probabilidad; Carnap lo llama probabilidad inductiva o lógica. Para mayores detalles, véase en particular R. Carnap, "On Inductive Logic", *Philosophy of Science*, vol. 12 (1945); *Logical Foundations of Probability* (Chicago, 1950; 2ª ed., 1962); *The Continuum of Inductive Methods* (Chicago, 1952); "The Aim of Inductive Logic", en E. Nagel, P. Suppes y A. Tarski, (comps.), *Logic, Methodology, and Philosophy of Science. Proceedings of the 1960 International Congress* (Stanford, 1962).

intentos hechos hasta ahora para formular una definición estadística precisa del grado de confirmación de una hipótesis parecen vulnerables a ciertas objeciones,[9] y varios autores [10] han expresado dudas acerca de la posibilidad de definir el grado de confirmación de una hipótesis como magnitud métrica, aunque algunos de ellos consideran posible, en ciertas condiciones, resolver al menos el problema —menos exigente— B(2), es decir, establecer normas de comparación no métrica entre hipótesis con respecto al grado de su confirmación. Una comparación apropiada de este tipo debería tomar en consideración una cantidad de factores diferentes; [11] pero, nuevamente, el número de los casos confirmatorios y el de los desconfirmatorios que contengan los elementos de juicio dados figuraría entre los más importantes de esos factores.

Así, de los dos problemas, A y B, A parece ser el básico; primero, porque no presupone la posibilidad de definir grados numéricos de confirmación o de comparar diferentes hipótesis con respecto a la extensión de su confirmación; y segundo, porque nuestras consideraciones indican que todo intento para resolver el problema B —a menos que permanezca en la etapa de un sistema axiomatizado sin interpretación— probablemente exija una definición precisa de los conceptos de caso confirmatorio y caso desconfirmatorio de una hipótesis antes de que se pueda proceder a definir grados numéricos de confirmación, o a establecer normas no métricas de comparación.

d) Ahora resulta claro que un análisis de la confirmación es de fundamental importancia también para el estudio de un problema central de la epistemología, a saber, el de la elaboración de normas de creencia racional o de criterios de asertibilidad garantizada. En la metodología de la ciencia empírica habitualmente se formula este problema como concerniente a las reglas que gobiernan la prueba de ensayo (*test*) y la ulterior aceptación o rechazo de hipótesis empíricas sobre la base de hallazgos experimentales u observacionales, mientras que en su versión epistemológica se suele formular el problema como concerniente a la convalidación de creencias con referencia a las percepciones, los datos sensoriales, etcétera. Pero de cualquier modo que se conciban los elementos de juicio finales, y sean cuales fueren los términos en los que se exprese esta concepción, el problema teórico es el mismo: caracterizar, en términos precisos y generales, las condiciones en las cuales puede decirse que un conjunto de elementos de juicio confirman o desconfirman una hipótesis de carácter empírico; y esto equivale, nuevamente, a nuestro problema A.

[9] Véase Karl Popper, *Logik der Forschung* (Viena, 1935), sección 80; Ernest Nagel, *l. c.*, sección 8, y "Probability and the Theory of Knowledge", *Philosophy of Science*, vol. 6 (1939); C. G. Hempel, "Le problème de la vérité", *Theoria* (Gotemburgo), vol. 3 (1937), sección 5, y "On the Logical Form of Probability Statements", *Erkenntnis*, vol. 7 (1937-1938), en particular sección 5. Véase también Morton White, "Probability and Confirmation", *The Journal of Philosophy*, vol. 36 (1939).

[10] Véase, por ejemplo, J. M. Keynes, *A Treatise on Probability* (Londres, 1929), en particular cap. III; Ernest Nagel, *Principles of the Theory of Probability*, en particular pág. 70; comparar también con el enunciado un poco menos definidamente escéptico de Carnap, en *l. c.* (nota 4) sección 3, pág. 427.

[11] Véase en particular el cuadro de tales factores que ofrece Ernest Nagel en *Principles of the Theory of Probability*, págs. 66-73.

e) El mismo problema surge cuando se intenta dar una enunciación precisa a los criterios empirista y operacionalista de la significación empírica de una oración (*sentence*); estos criterios, como es bien sabido, se formulan con referencia a la testabilidad teórica de la oración por medio de elementos de juicio experienciales,[12] y el concepto de testabilidad teórica, como señalamos antes, está íntimamente relacionado con los conceptos de confirmación y desconfirmación.[13]

Considerando la gran importancia del concepto de confirmación, es sorprendente que no se haya elaborado hasta ahora ninguna teoría sistemática de la relación no-cuantitativa de confirmación. Este hecho quizá refleje la suposición tácita de que los conceptos de confirmación y desconfirmación tienen un significado suficientemente claro como para hacer innecesarias, o relativamente triviales, las definiciones explícitas. Y en verdad, como veremos más adelante, hay ciertos aspectos asociados en general con la noción intuitiva de elemento de juicio confirmatorio que, en primera instancia, parecen muy apropiados para servir como características definitorias de la confirmación. Un examen más minucioso revelará que las definiciones que pueden obtenerse de este modo presentan serias deficiencias, y pondrá en claro que una definición adecuada de confirmación plantea considerables dificultades.

Ahora bien, la existencia misma de tales dificultades sugiere la pregunta acerca de si el problema que estamos considerando no reposa en una suposición falsa. Quizá no haya criterios objetivos de confirmación; quizá la decisión relativa a si una hipótesis dada es aceptable a la luz de un conjunto dado de elementos de juicio no pueda estar más sujeta a reglas racionales y objetivas que el proceso de inventar una hipótesis o teoría científica; quizás, en último análisis, sea una "sensación de evidencia" o un sentimiento de plausibilidad ante los datos atinentes a la cuestión lo que, en última instancia, decide si una hipótesis es científicamente aceptable.[14] Esta idea es comparable a la opinión de que la validez de una prueba matemática o de un razonamiento lógico debe ser juzgada, en última instancia, con referencia a un sentimiento de corrección o de convicción. Ambas tesis deben ser rechazadas por razones análogas, ya que ambas implican una confusión de consideraciones lógicas con consideraciones psicológicas. Evidentemente, la existencia o inexistencia de un sentimiento de convicción ante las pruebas de una aserción es una cuestión subjetiva que varía de una persona a otra, y en la misma persona en el transcurso del tiempo; a menudo es engañosa y no puede servir, por cierto, como condición necesaria ni

[12] Véase, por ejemplo, A. J. Ayer, *Language, Truth and Logic* (Londres y Nueva York, 1936), cap. I; R. Carnap, "Testability and Meaning", secciones 1, 2 y 3; H. Feigl, "Logical Empiricism" (en *Twentieth Century Philosophy*, comp. por Dagobert D. Runes, Nueva York, 1943); P. W. Bridgman, *The Logic of Modern Physics* (Nueva York, 1928).

[13] Debe observarse, sin embargo, que en su ensayo "Testability and Meaning", R. Carnap ha elaborado definiciones de la testabilidad y la confirmabilidad que evitan toda referencia a los conceptos de confirmación y desconfirmación. En verdad, en este estudio no se hace propuesta alguna para definir esos conceptos.

[14] Una concepción de este tipo ha sido expresada, por ejemplo, por M. Mandelbaum, en "Causal Analyses in History", *Journal of the History of Ideas*, vol. 3 (1942); véase en particular la sección 8.

suficiente para determinar la corrección de una aserción determinada.[15] Una reconstrucción racional de las normas de la convalidación científica, pues, no puede implicar referencia alguna a una sensación de evidencia, sino que debe basarse en criterios objetivos. De hecho, parece razonable exigir que los criterios de confirmación empírica, además de tener un carácter objetivo, no contengan ninguna referencia al objeto específico de la hipótesis o de los elementos de juicio en cuestión; debe ser posible, creemos, establecer criterios puramente formales de confirmación, análogamente a la manera como la lógica deductiva suministra criterios puramente formales para determinar la validez de la inferencia deductiva.

Con este objetivo en vista, pasaremos ahora a estudiar el concepto no cuantitativo de confirmación. Comenzaremos examinando algunas concepciones corrientes de la confirmación y poniendo de manifiesto sus defectos lógicos y metodológicos. En el curso de este análisis elaboraremos un conjunto de condiciones para establecer el valor de cualquier definición propuesta de la confirmación. Finalmente, construiremos una definición de confirmación que satisfaga todas esas normas generales.

3. *El criterio de confirmación de Nicod y sus inconvenientes*

Consideraremos primero una concepción de la confirmación que subyace en muchos de los estudios recientes sobre la inducción y el método científico. Jean Nicod hizo una enunciación muy explícita de esta concepción en el siguiente pasaje: "Consideremos la fórmula de la ley: *A implica B*. ¿De qué manera puede influir sobre su probabilidad una proposición particular, o más brevemente, un hecho? Si este hecho consiste en la presencia de B en un caso de A, es favorable a la ley "*A implica B*"; por el contrario, si consiste en la ausencia de B en un caso de A, es desfavorable a ella. Es concebible que tengamos aquí los únicos dos modos directos en los cuales un hecho puede influir sobre la probabilidad de una ley... De este modo, la influencia total de las verdades o hechos particulares sobre la probabilidad de proposiciones o leyes universales actuaría por medio de estas dos relaciones elementales que llamaremos *confirmación* e *invalidación*".[16] Obsérvese que la aplicabilidad de este criterio está restringida a hipótesis de la forma "*A implica B*". Toda hipótesis *H* de este tipo puede expresarse en la notación de la lógica simbólica[17] por medio de una oración condicional universal tal como, en el caso más simple:

$$(x)[P(x) \supset Q(x)]$$

esto es, "para todo objeto *x*, si *x* es un *P*, entonces *x* es un *Q*"; o también,

[15] Véase en particular la declaración de Popper, *loc. cit.*, sección 8.

[16] Jean Nicod, *Foundations of Geometry and Induction* (trad. al inglés por P. P. Wiener) (Londres, 1930, pág. 219). Véase también el examen de R. M. Eaton "Confirmation and Information", que se basa en las concepciones de Nicod; está incluido en el cap. III de su *General Logic* (Nueva York, 1931).

[17] En este ensayo sólo usamos las formas más elementales de esta notación. El simbolismo es esencialmente el de *Principia Mathematica*, excepto que usamos paréntesis en lugar de puntos, y simbolizamos el cuantificador existencial por "(E)" en lugar de usar la "E" invertida.

"la aparición de la cualidad P implica la aparición de la cualidad Q". Según el criterio anterior, esta hipótesis está confirmada por un objeto a, si a es P y Q, y la hipótesis está confirmada por a, si a es P pero no Q.[18] En otras palabras, un objeto confirma una hipótesis condicional universal si y sólo si satisface tanto el antecedente (aquí: "$P(x)$") como el consecuente (aquí: "$Q(x)$") del condicional; desconfirma la hipótesis si y sólo si satisface el antecedente pero no el consecuente del condicional; y (agregamos esto al enunciado de Nicod) es neutral o ajeno a la hipótesis, si no satisface el antecedente.

Puede extenderse fácilmente este criterio de modo que sea también aplicable a condicionales universales que contengan más de un cuantificador, por ejemplo, "los gemelos siempre se parecen" o, en notación simbólica, "(x) (y) (Gemelos (x,y) $\supset$ Par. (x,y))". En estas circunstancias, un caso confirmatorio consiste en un par, un trío, etc., ordenado de objetos que satisfacen el antecedente y el consecuente del condicional. (En el último ejemplo, dos personas cualesquiera que fueran gemelas y se parecieran confirmarían la hipótesis; dos gemelos que no se parecieran, las desconfirmarían; y dos personas cualesquiera que no fueran gemelas, se parecieran o no, constituirían elementos de juicio ajenos a la cuestión.)

Al criterio mencionado lo llamaremos criterio de Nicod.[19] El mismo enuncia de modo explícito algo que constituye, quizá, la interpretación tácita más común del concepto de confirmación. Aunque en apariencia es muy adecuado, presenta serios inconvenientes, como mostraremos a continuación.

a) Primero, la aplicabilidad de este criterio está restringido a hipótesis de la forma condicional universal; no brinda ninguna norma de confirmación para las hipótesis existenciales (tales como "existe vida orgánica en otras estrellas", o "la poliomielitis es causada por un virus") ni para hipótesis cuya formulación explícita exige el uso de cuantificadores universales y existenciales (tales como "todo ser humano muere al cabo de un número finito de años después de su nacimiento", o como la hipótesis psicológica "podéis burlar a todo el mundo durante algún tiempo y a algunas personas constantemente, pero no podéis burlar a todo el mundo constantemente", que puede ser simbolizado por "$(x)(Et)Br(x,t)$ · $(Ex)(t)Br(x,t)$ · $\sim$ (x) (t) $(Br(x,t)$", donde "$Br(x,t)$" representa a "podéis burlar a una persona x en el tiempo t"). Destaquemos, pues, el desiderátum de establecer un criterio de confirmación que sea aplicable a hipótesis de *cualquier* forma.[20]

[18] (Agregada en 1964.) Más precisamente, en el lenguaje de Nicod deberíamos decir que la hipótesis queda confirmada por la proposición de que a es P y Q, y desconfirmada por la proposición de que a es P pero no Q.

[19] Elegimos este término por conveniencia y teniendo en cuenta la anterior formulación explícita dada por Nicod; por supuesto, ello no implica que esta concepción de la confirmación tenga su origen en Nicod.

[20] Para una formulación rigurosa del problema, primero es necesario establecer suposiciones acerca de los medios de expresión y la estructura lógica del lenguaje en el cual se supone formuladas las hipótesis; el desiderátum, pues, exige una definición de confirmación que sea aplicable a cualquier hipótesis que pueda expresarse en el lenguaje dado. Hablando en general, el problema se hace cada vez más difícil a medida que aumenta la riqueza y la complejidad del lenguaje de la ciencia adoptado.

b) Pasemos ahora a un segundo inconveniente del criterio de Nicod. Consideremos las dos oraciones:

O_1: "(x) [Cuervo(x) $\supset$ Negro(x)]";

O_2: "(x) [$\sim$ Negro(x) $\supset$ $\sim$ Cuervo(x)]"

(es decir, "todos los cuervos son negros" y "todo lo que no es negro no es un cuervo"), y sean *a, b, c, d* cuatro objetos tales que *a* es un cuervo y es negro, *b* es un cuervo pero no es negro, *c* no es un cuervo pero es negro, y *d* no es un cuervo ni es negro. Entonces, de acuerdo con el criterio de Nicod, *a* confirmaría O_1 pero sería neutral con respecto a O_2; *b* desconfirmaría O_1 y O_2; *c* sería neutral con respecto a O_1 y O_2; y *d* confirmaría O_2 pero sería neutral con respecto a O_1.

Pero Q_1 y Q_2 son lógicamente equivalentes: tienen el mismo contenido, son formulaciones diferentes de la misma hipótesis. Sin embargo, por el criterio de Nicod, los objetivos *a* y *d* confirmarían una de las dos oraciones, pero serían neutrales con respecto a la otra. Esto significa que el criterio de Nicod hace depender la confirmación, no solamente del contenido de la hipótesis, sino también de su formulación.[21]

Una consecuencia notable de esta situación es que toda hipótesis a la cual sea aplicable el criterio —es decir, todo condicional universal— puede ser formulada en una forma para la cual no pueden existir casos confirmatorios. Así, por ejemplo, la oración:

$$(x)\,[(\text{Cuervo}(x) \cdot \sim \text{Negro}(x)) \supset (\text{Cuervo}(x) \cdot \sim \text{Cuervo}(x))]$$

puede reconocerse fácilmente como equivalente a O_1 y O_2; sin embargo, ningún objeto puede confirmar esta oración, es decir, satisfacer su antecedente y su consecuente, pues éste es contradictorio. Una transformación análoga es aplicable, por supuesto, a cualquier otra oración de forma condicional universal.

4. *La condición de equivalencia*

Los resultados que acabamos de obtener llaman la atención a la siguiente condición que debe satisfacer un concepto adecuadamente definido de confirmación, y a la luz del cual debe rechazarse el criterio de Nicod:

Condición de equivalencia: Todo lo que confirme (desconfirme) a una de dos oraciones equivalentes, también confirma (desconfirma) a la otra.

El cumplimiento de esta condición independiza la confirmación de una hipótesis de la manera de formularla, y sin duda se admitirá que es una condición necesaria para la adecuación de cualquier criterio de confirmación que se proponga. De lo contrario, la cuestión de si ciertos datos confirman una hipótesis dada debería responderse diciendo: "Eso depende de cuál de las diferentes formulaciones de la hipótesis se considere", lo cual parece absurdo. Además, y éste es un punto más importante que la apelación a una sensación de absurdo, una definición adecuada de confirmación

[21] Esta dificultad fue señalada, en esencia, en mi artículo "Le problème de la vérité", *Theoria* (Gotemburgo), vol. 3 (1937), en particular pág. 222.

deberá hacer justicia a la manera como las hipótesis empíricas funcionan en contextos teóricos científicos, en cuanto explicaciones y predicciones; pero cuando se usan las hipótesis para los fines de la explicación o la predicción,[22] actúan como premisas de un razonamiento deductivo cuya conclusión es una descripción del suceso que se quiere explicar o predecir. La deducción está regida por los principios de la lógica formal y, según ésta, una deducción válida lo sigue siendo si se reemplazan algunas o todas las premisas por enunciados diferentes pero equivalentes. Ahora bien, todo científico se sentirá libre, en cualquier razonamiento teórico que incluya ciertas hipótesis, para usar éstas en cualquiera de las formulaciones equivalentes que sean más convenientes para el desarrollo de sus conclusiones. Pero si adoptamos un concepto de confirmación que no satisfaga la condición de equivalencia, entonces sería posible, y hasta necesario, argüir en ciertos casos que es un procedimiento científico correcto basar una predicción en una hipótesis dada si se la formula en una oración O_1; pero que es totalmente inadmisible basar la predicción (digamos, para conveniencia de la deducción en una formulación equivalente O_2, porque no se dispone de ningún elemento de juicio confirmatorio para O_2. Así, la condición de equivalencia debe ser considerada como una condición necesaria para la adecuación de cualquier definición de confirmación.

5. *Las paradojas de la confirmación*

Quizá parezca que estamos insistiendo en lo obvio al destacar la necesidad de la condición de equivalencia. Es probable que esta impresión desaparezca si consideramos ciertas consecuencias que derivan de una combinación de la condición de equivalencia con una suposición muy natural y plausible concerniente a una condición suficiente para la confirmación.

La esencia de la crítica que hemos dirigido contra el criterio de Nicod es que no puede servir, por cierto, como condición necesaria de la confirmación. Así, en el ejemplo dado al comienzo de la sección 3, el objeto *a* confirma O_1 y, por ende, debe considerarse que también confirma O_2, lo cual no sucede de acuerdo con el criterio de Nicod. La satisfacción de este criterio, pues, no es una condición necesaria para que los elementos de juicio sean confirmatorios.

Por otra parte, puede considerarse, con todo, que el criterio de Nicod enuncia una condición suficiente particularmente obvia e importante de la confirmación. En verdad, si nos limitamos a las hipótesis condicionales universales de una variable,[23] como O_1 y O_2 en el ejemplo anterior, parece per-

[22] Para una descripción más detallada de la estructura lógica de la explicación y la predicción científicas, véase C. G. Hempel, "The Function of General Laws in History", *The Journal of Philosophy*, vol. 39 (1942), especialmente secciones 2, 3 y 4. La caracterización, dada en este capítulo y en el texto anterior, de las explicaciones y predicciones como razonamientos de estructura lógica deductiva es una simplificación: como veremos en la sección 7 de este ensayo, las explicaciones y las predicciones a menudo incluyen pasos "casi-inductivos", además de los deductivos. Este punto, sin embargo, no afecta a la validez del argumento anterior.

[23] Esta restricción es esencial: en su forma general, que se aplica a condicio-

fectamente razonable calificar un objeto como confirmatorio de una hipótesis si satisface su antecedente y su consecuente. La plausibilidad de esta opinión será corroborada en el curso de nuestros análisis posteriores.

Luego, convendremos en que si *a* es un cuervo y es negro, entonces confirma ciertamente O_1: "(x) (Cuervo $(x) \supset$ (Negro(x)))" y si *d* no es un cuervo ni es negro, ciertamente confirma O_2: "$(x)[\sim$ Negro $(x) \supset \sim$ Cuervo(x)]".

Combinemos ahora esta estipulación simple con la condición de equivalencia. Puesto que O_1 y O_2 son equivalentes, *d* también confirma O_1; de este modo, debemos reconocer que todo objeto que no sea negro ni sea un cuervo confirma O_1. Por consiguiente, todo lápiz rojo, toda hoja verde, toda vaca amarilla, etc., se convierte en elemento de juicio confirmatorio de la hipótesis de que todos los cuervos son negros. Esta sorprendente consecuencia de dos suposiciones muy apropiadas (la condición de equivalencia y la anterior condición suficiente de la confirmación) puede ampliarse aún: es posible mostrar fácilmente que la oración O_1 es equivalente a O_3: "(x) [(Cuervo(x) v $\sim$(Cuervo(x))) $\supset$ ($\sim$Cuervo(x) v Negro(x))]", es decir, "todo lo que es o no es un cuervo, o no es un cuervo o es negro". De acuerdo con la anterior condición suficiente, O_3 está confirmada, ciertamente, por todo objeto, llamémosle *e*, tal que (1) *e* es o no es un cuervo, y además, (2) *e* no es un cuervo o es también negro. Puesto que (1) es analítica, estas condiciones se reducen a (2). En virtud de la condición de equivalencia, pues, debemos considerar como confirmatorio de O_1 todo objeto que no sea un cuervo o sea también negro (en otras palabras, todo objeto que no sea en absoluto un cuervo o sea un cuervo negro).

De los cuatro objetos caracterizados en la sección 3, *a*, *c* y *d* constituirían, pues, elementos de juicio confirmatorios de O_1, mientras que *b* la des-

---

nales universales de cualquier número de variables, el criterio de Nicod ni siquiera puede ser concebido como una condición suficiente de la confirmación. Se verá esto mediante el siguiente ejemplo, que es bastante sorprendente:

Consideremos la hipótesis:

$$O_1 : (x) \; (y) \; [\sim (R\,(x,y) \cdot R\,(y,x)\,) \supset (R\,(x,y) \cdot \sim R\,(y,x)\,)].$$

Sean *a* y *b* dos objetos tales que $R(a,b)$ y $\sim R(b,a)$. Entonces, evidentemente, el par (a,b) satisface el antecedente y el consecuente del condicional universal O_1. Luego, si se admite que el criterio de Nicod en su forma general enuncia una condición suficiente para la confirmación, (a,b) constituye un elemento de juicio confirmatorio de O_1. Pero puede demostrarse que O_1 es equivalente a:

$$O_2 : (x) \; (y) \; R\,(x,y).$$

Ahora bien, por hipótesis, tenemos $\sim R(b,a)$, lo cual contradice de plano a O_2 y, por ende, a O_1. Así, el par (a,b), si bien satisface el antecedente y el consecuente del condicional universal O_1, en realidad constituye un elemento de juicio desconfirmatorio del tipo más fuerte (elemento de juicio desconfirmatorio concluyente, como diremos más adelante) de esa oración. Este ejemplo revela una debilidad sorprendente y —según mi conocimiento— hasta ahora inadvertida de la concepción de la confirmación que subyace en el criterio de Nicod. Para comprender la relación de nuestro ejemplo con la formulación original de Nicod, sean *A* y *B*: $\sim(R(x,y) \cdot R(y,x))$ y $R(x,y) \cdot \sim R(y,x)$, respectivamente. Entonces, O_1 afirma que *A* implica *B*, y el par (a,b) es un caso de la presencia de *B* en presencia de *A*; según el criterio de Nicod, esto es favorable a O_1.

confirmaría. Esto implica que todo no-cuervo representa un elemento de juicio confirmatorio de la hipótesis de que todos los cuervos son negros.[24]

Llamaremos a estas implicaciones de la condición de equivalencia y de la mencionada condición suficiente de la confirmación las *paradojas de la confirmación.*

¿Cómo deben abordarse estas paradojas? Renunciar a la condición de equivalencia no sería una solución aceptable, como lo revelan las consideraciones expuestas en la sección 4. Tampoco parece posible prescindir de la estipulación según la cual un objeto que satisfaga dos condiciones, C_1 y C_2, debe ser considerado como confirmante de una hipótesis general en el sentido de que todo objeto que satisfaga C_1 también satisfará C_2.

Pero la deducción de los anteriores resultados paradójicos reposa en otra suposición que habitualmente se halla tácita, a saber, que el significado de las hipótesis empíricas generales, como la de que todos los cuervos son negros, o la de que todas las sales de sodio dan un color amarillo al entrar en combustión, pueden ser adecuadamente expresadas por medio de oraciones de forma condicional universal, por ejemplo, "(x) [Cuervo(x) $\supset$ $\supset$ Negro(x)]" y "(x) (Sod. Sal(x) $\supset$ Comb. Amar.(x))", etcétera. Quizá sea necesario modificar este modo habitual de presentación. ¿Eliminaría tal modificación, automáticamente, las paradojas de la confirmación? Si no es así, sólo parece quedar una alternativa, a saber, mostrar que el aparente carácter paradójico de esas consecuencias se debe a un malentendido y puede ser eliminado de modo que no subsistan dificultades teóricas. Consideremos ahora estas dos posibilidades por turno. Las subsecciones 5.1.1 y 5.1.2 están dedicadas al examen de dos propuestas diferentes de representación modificada de las hipótesis generales; en la subsección 5.2 examinaremos la segunda alternativa, esto es, la posibilidad de reducir la apariencia paradójica a un malentendido.

5.1.1. Se ha señalado a menudo que la lógica aristotélica, de acuerdo con el uso cotidiano prevaleciente, asigna contenido existencial a las oraciones de la forma "todos los P son Q", mientras que para la lógica moderna una oración condicional universal carece de todo contenido existencial. Así, la oración:

"(x) [Sirena(x) $\supset$ Verde(x)]"

no implica la existencia de sirenas: simplemente afirma que todo objeto no es en absoluto una sirena, o es una sirena verde: es verdadera simplemente por el hecho de que no hay sirenas. Pero las leyes y las hipótesis generales de la ciencia, podría argüirse, pretenden tener contenido existencial, y se podría tratar de expresar este contenido completando el condicional univer-

[24] (Agregada en 1964.) Pueden observarse las siguientes consecuencias "paradójicas" adicionales de nuestras dos condiciones: Toda hipótesis de forma condicional universal puede ser reformulada de manera equivalente como otra hipótesis de la misma forma que, aunque sea verdadera, no puede en absoluto tener casos confirmatorios en el sentido de Nicod, pues la proposición según la cual un objeto dado satisface el antecedente y el consecuente de la segunda hipótesis es contradictoria. Por ejemplo, "$(x)[P(x) \supset Q(x)]$" es equivalente a la oración "$(x)[(P(x) \cdot \sim Q(x)) \supset (P(x) \cdot \cdot \sim P(x))]$", cuyo consecuente no es verdadero en ningún caso.

sal corriente con una cláusula existencial. Por ejemplo, la hipótesis de que todos los cuervos son negros se expresaría por medio de la oración O_1: "$[(x)(\text{Cuervo}(x) \supset \text{Negro}(x))] \cdot (Ex)\text{Cuervo}(x)$", y la hipótesis de que ninguna cosa no negra es un cuervo, por O_2: "$(x)[\sim \text{Negro}(x) \supset \sim \text{Cuervo}(x)] \cdot (Ex) \sim \text{Negro}(x)$". Es indudable que estas oraciones no son equivalentes, y de los cuatro objetos caracterizados en la sección 3, parte (*b*), sólo de *a* podría decirse razonablemente que confirma O_1, y sólo *d* confirma O_2. Sin embargo, este método de evitar las paradojas de la confirmación es vulnerable a serias objeciones.

a) Ante todo, la representación de toda hipótesis general por la conjunción de un condicional universal y una oración existencial invalidaría muchas inferencias lógicas generalmente consideradas admisibles en un razonamiento teórico. Así, por ejemplo, las afirmaciones de que todas las sales de sodio dan un color amarillo al entrar en combustión y de que todo lo que no dé tal color amarillo en la combustión no es una sal de sodio son lógicamente equivalentes, según la interpretación y el uso habituales, y su representación mediante condicionales universales mantiene esta equivalencia. Pero si se agregan cláusulas existenciales, las dos afirmaciones ya no serán equivalentes, como lo ejemplificó antes el caso análogo de O_1 y O_2.

b) En segundo término, la formulación habitual de las hipótesis generales en las ciencias empíricas no contiene cláusulas existenciales, ni por lo general determina, ni siquiera indirectamente, tal cláusula de manera no ambigua. Por ejemplo, consideremos la hipótesis de que si una persona, después de recibir una inyección de una sustancia que se quiere ensayar, presenta una reacción positiva en la piel, entonces tiene difteria. ¿Debemos entender aquí que la cláusula existencial se refiere a personas en general, a personas que reciben la inyección o a personas que después de recibir la inyección presentan una reacción positiva en la piel? Se hace necesario tomar una decisión más o menos arbitraria, y cada una de las decisiones posibles da una interpretación diferente de la hipótesis, aunque ninguna de ellas parece realmente implicada por ésta.

c) Finalmente, de muchas hipótesis universales no puede decirse que impliquen en modo alguno una cláusula existencial. Así, puede suceder que de determinada teoría astrofísica se deduzca una hipótesis universal concerniente al carácter de los fenómenos que se producirían en ciertas condiciones específicas extremas. Una hipótesis de este tipo no necesita implicar (y, por lo general, no implica) que se realicen alguna vez tales condiciones extremas; no tiene contenido existencial alguno. O bien, consideremos una hipótesis biológica que afirma que, cuando se cruzan el hombre y el mono, los descendientes tienen tales o cuales características. Se trata de una hipótesis general; se la podría considerar como una mera conjetura o como una consecuencia de una teoría genética más amplia, de la cual pueden haberse puesto a prueba ya con resultados positivos otras consecuencias. Pero es indudable que la hipótesis no implica una cláusula existencial por la cual el tipo de hibridación aludida realmente se produzca en algún momento.

5.1.2. Quizá pueda decirse que el aparente carácter paradójico de los casos examinados al comienzo de la sección 5 surge de la sensación de que

la hipótesis según la cual todos los cuervos son negros se refiere a cuervos, y no a cosas no negras o a todas las cosas. El uso de una cláusula existencial fue un intento para poner de manifiesto esta presunta pecularidad de la hipótesis. El intento ha fracasado: si deseamos expresar el punto en cuestión, deberemos buscar un recurso más efectivo. Surge naturalmente la idea de representar una hipótesis general mediante el usual condicional universal, complementado con la indicación del "campo de aplicación" específico de la hipótesis: así, podríamos representar la hipótesis de que todos los cuervos son negros por la oración "(x) [Cuervo $(x) \supset$ Negro (x)]" o cualquiera de sus equivalentes, más la indicación "Clase de cuervos", que caracteriza el campo de aplicación: y luego podríamos exigir que todo caso confirmatorio pertenezca al campo de aplicación. Este procedimiento excluiría los objetos *c* y *d* de los que constituyen elementos de juicio confirmatorios y, de este modo, evitaría las consecuencias indeseables del recurso de la cláusula existencial que señalamos en 5.1.1 (*c*). Pero, aparte de esta ventaja, el segundo método es vulnerable a objeciones similares a las que pueden dirigirse contra el primero: (*a*) La manera como se usan las hipótesis generales en la ciencia nunca supone la determinación de un campo de aplicación; y la elección de este último en una formulación simbólica de una hipótesis determinada introduce, pues, una considerable parte de arbitrariedad. Con respecto a una hipótesis científica que diga que todos los *P* son *Q*, no puede decirse simplemente que el campo de aplicación sea la clase de todos los *P*; pues una hipótesis como la de que todas las sales de sodio dan un color amarillo al entrar en combustión halla importantes aplicaciones en ensayos de resultados negativos: por ejemplo, puede aplicarse a una sustancia de la que no se sabe si contiene o no sales de sodio, o si da o no color amarillo al entrar en combustión; y si la llama no da un color amarillo, la hipótesis sirve para establecer la ausencia de sales de sodio. Lo mismo es verdadero de todas las otras hipótesis utilizadas para ensayos de este tipo. (*b*) También, la utilización constante de un campo de aplicación en la formulación de hipótesis generales traería aparejadas considerables dificultades lógicas, y no tendría parangón alguno en los procedimientos teóricos de la ciencia, donde se someten las hipótesis a diversos tipos de transformaciones e inferencias lógicas sin atender para nada a cambios en los campos de aplicación. Este método de resolver las paradojas, pues, equivaldría a eludir el problema por medio de un recurso *ad hoc* que no halla justificación en el procedimiento científico real.

5.2. Hemos examinado dos alternativas al método habitual de representar las hipótesis generales mediante condicionales universales, y ninguna de ellas resultó ser un medio adecuado para eliminar las paradojas de la confirmación. Trataremos ahora de mostrar que lo erróneo no reside en el método habitual de construir y representar hipótesis generales, sino en confiar en una engañosa intuición en la materia: la impresión de que se trata de una situación paradójica no tiene fundamento objetivo, sino que es una ilusión psicológica.

a) Una fuente de malentendidos es la idea, a la que ya aludimos, de que una hipótesis de la forma simple "todo *P* es un *Q*", como "todas las sales de sodio dan un color amarillo al quemarse", afirma algo acerca de

determinada clase limitada de objetos, a saber, la clase de todos los *P*, solamente. Esta idea supone una confusión entre consideraciones lógicas y consideraciones prácticas: nuestro interés por la hipótesis puede concentrarse en su aplicabilidad a esta clase particular de objetos, pero la hipótesis afirma algo de *todos* los objetos (dentro del tipo lógico de la variable que aparece en la hipótesis, el cual, en el caso de nuestro último ejemplo, *podría* ser la clase de todos los objetos físicos) y, en verdad, también impone restricciones a todos ellos. En realidad, una hipótesis de la forma "todo *P* es un *Q*" prohibe la aparición de yodo que tenga la propiedad *P* pero carezca de la propiedad *Q*; es decir, restringe todos los objetos, sean cuales fueren, a la clase de aquellos que o bien carecen de la propiedad *P*, o bien tienen también la propiedad *Q*. Pero todo objeto pertenece a esta clase o está fuera de ella, y por ende, todo objeto —no solamente los *P*— se ajusta a la hipótesis o la viola; no hay objeto al que no se aluda implícitamente en una hipótesis de este tipo. En particular, todo objeto que no es una sal de sodio o da un color amarillo al quemarse se ajusta a ella, con lo cual sustenta la hipótesis de que todas las sales de sodio dan un color amarillo al entrar en combustión; todo otro objeto se opone a esta hipótesis.

La debilidad de la idea considerada se pone de manifiesto también por la observación de que la clase de objetos acerca de los cuales se supone que una hipótesis afirma algo no está en modo alguno claramente determinada y cambia según el contexto, como mostramos en 5.1.2. (*a*).

b) Una segunda fuente importante de paradojas en ciertos casos de confirmación surge de la siguiente consideración.

Supongamos que en apoyo de la afirmación "todas las sales de sodio dan un color amarillo al entrar en combustión" alguien adujera un experimento en el cual un trozo de hielo puro fuera sostenido en una llama incolora y no diera coloración amarilla a la llama. Este resultado confirmaría la aserción "todo lo que no da color amarillo al quemarse no es una sal de sodio" y, por consiguiente, en virtud de la condición de equivalencia, confirmaría la formulación original. ¿Por qué nos da esto la impresión de ser paradójico? La razón de ello se aclara cuando comparamos la situación anterior con el caso en el cual colocamos un objeto cuya composición química aún se deconoce en una llama y no da coloración amarilla, a la par que el análisis ulterior revela que no contiene sales de sodio. Este resultado, admitiríamos sin duda, es el que cabía esperar sobre la base de la hipótesis de que todas las sales de sodio dan un color amarillo al quemarse, en cualquiera de sus diversas formulaciones equivalentes en que la expresemos; de este modo, los datos obtenidos constituyen elementos de juicio confirmatorios de la hipótesis. Ahora bien, la única diferencia entre las dos situaciones consideradas es que en el primer caso se nos dice de antemano que la sustancia de prueba es hielo, y sucede que "sabemos de algún modo" que el hielo no contiene sales de sodio; la consecuencia de esto es que el resultado del test de color de la llama es totalmente ajeno a la confirmación de la hipótesis y, por ende, no puede constituir un elemento de juicio nuevo para nosotros. En verdad, si la llama no se pone amarilla, la hipótesis exige que la sustancia no contenga sales de sodio, y sabemos de antemano que el hielo no las contiene; y si la llama se pone amarilla, la

hipótesis no impondría ninguna restricción adicional a la sustancia: luego, cualquiera de los resultados posibles de experimento estaría de acuerdo con la hipótesis.

El análisis de este ejemplo aclara un punto de carácter general: en los casos aparentemente paradójicos de confirmación, no juzgamos en realidad la relación del elemento de juicio dado *E* solamente con la hipótesis *H* (pasamos por alto la ficción metodológica, característica de todo caso de confirmación, de que no tenemos elementos de juicio atinentes a *H* que no sean los incluidos en *E*), sino que introducimos tácitamente una comparación de *H* con un conjunto de elementos de juicio formado por *E* en conjunción con la información adicional de que disponemos. En nuestro ejemplo, esta información incluye el conocimiento: (1) de que la sustancia usada en el experimento es hielo, y (2) de que el hielo no contiene sales de sodio. Si suponemos dada esta información adicional, entonces, por supuesto, el resultado del experimento no puede agregar fuerza alguna a la hipótesis en consideración. Pero si tomamos la precaución de evitar esta referencia tácita a un conocimiento adicional (lo cual cambia por completo el carácter del problema) y si formulamos la cuestión relativa al carácter confirmatorio de los elementos de juicio de una manera adecuada al concepto de confirmación utilizado en este artículo, debemos preguntar: dado un objeto *a* (que es un trozo de hielo, pero no se incluye este hecho en los elementos de juicio) y dado el hecho de que *a* no pone amarilla la llama y no es una sal de sodio, ¿constituye *a* un elemento de juicio confirmatorio de la hipótesis? Y ahora, sea *a* hielo o cualquier otra sustancia, es claro que la respuesta debe ser afirmativa y la paradoja desaparece.

Hasta ahora, en la sección (*b*), hemos considerado principalmente el tipo de paradoja ejemplificada por la aserción de que todo no-cuervo no-negro constituye un elemento de juicio confirmatorio de la hipótesis "todos los cuervos son negros". Sin embargo, la idea general que acabamos de esbozar se aplica también a los casos aun más extremos ejemplificados por la aserción de que todo no-cuervo y todo objeto negro confirma la hipótesis en cuestión. Ejemplifiquemos esto con respecto al último caso. Si los elementos de juicio dados, *E* —esto es, en el sentido de la ficción metodológica exigida, todos los datos atinentes a la hipótesis—, consisten solamente en un objeto que, además, es negro, entonces puede decirse razonablemente que *E* da apoyo aun a la hipótesis de que todos los objetos son negros, y *a fortiori* *E* da apoyo a la aserción más débil de que todos los cuervos son negros. En este caso, nuevamente, nuestro conocimiento de que no todos los objetos son negros tiende a crear una impresión de paradoja que no tiene justificación lógica. Pueden tratarse de manera análoga otros casos paradójicos de confirmación. De este modo resulta que las paradojas de la confirmación, tales como las formulamos antes, se deben a una descaminada intuición en la cuestión, no a una falla lógica en las dos estipulaciones de las cuales derivan.[25, 26]

[25] La idea básica de la sección (*b*) del anterior análisis se debe al doctor Nelson Goodman, a quien quiero reiterar mi agradecimiento por la ayuda que me brindó, a través de muchas discusiones, para aclarar mis ideas sobre este punto.

[26] Las consideraciones presentadas en la sección (*b*), aunque no tienen el

6. *La confirmación concebida como una relación entre oraciones*

Nuestro análisis del criterio de Nicod nos ha conducido hasta ahora a dos resultados principales: al rechazo de este criterio por sus diversas deficiencias y a la emergencia de la condición de equivalencia como condición necesaria de adecuación para toda definición de la confirmación que se proponga. Debemos considerar ahora otro aspecto del criterio de Nicod. En nuestra formulación de este criterio concebimos la confirmación como una relación diádica entre un objeto o un conjunto ordenado de objetos, que representan los elementos de juicio, y una oración, que representa la

---

mismo contenido, reflejan también la influencia de los aclaradores análisis de las paradojas realizados por la metodóloga y lógica polaca Janina Hosiasson-Lindenbaum; véase su artículo "On Confirmation", *The Journal of Symbolic Logic*, vol. 5 (1940), particularmente la sección 4. Atrajeron la atención de la doctora Hosiasson hacia las paradojas mi artículo "Le problème de la vérité" y discusiones sostenidas conmigo. Por lo que sé, la suya es hasta ahora la única publicación que presenta un intento explícito por resolver el problema. Su solución se basa en una teoría de los grados de confirmación desarrollada en la forma de un sistema axiomático no interpretado, y la mayoría de sus argumentos presuponen esta armazón teórica. Sin embargo, he aprovechado algunas de las observaciones más generales de la señorita Hosiasson que guardan relación con el análisis de las paradojas del concepto no graduado o cualitativo de confirmación que constituye el objeto del presente estudio.

Un punto de los comentarios de Hosiasson que se basa en su teoría de los grados de confirmación presenta particular interés, por lo que quiero examinarlo brevemente. Enunciado con referencia a la hipótesis de los cuervos, consiste en la sugerencia de que el hallazgo de un objeto no negro que no sea un cuervo, si bien constituye un elemento de juicio confirmatorio de la hipótesis, aumentaría el grado de confirmación de la hipótesis en un grado menor que el hallazgo de un cuervo que sea negro. Se dice que esto es así porque la clase de los cuervos es mucho menos numerosa que la de todos los objetos no negros, de modo que —para dar a la idea una forma sugerente aunque un tanto engañosa— el hallazgo de un cuervo negro confirma una parte mayor del contenido total de la hipótesis que el hallazgo de un no-cuervo no-negro. De hecho, a partir de las suposiciones básicas de su teoría, la señorita Hosiasson logra deducir un teorema según el cual el anterior enunciado acerca del aumento relativo en el grado de confirmación será válido siempre que realmente el número de todos los cuervos sea pequeño comparado con el número de todos los objetos no-negros. Pero, ¿tiene fundamento esta suposición numérica en el caso presente y, análogamente, en todos los otros casos "paradójicos"? La respuesta depende, en parte, de la estructura lógica del lenguaje de la ciencia. Si se usa un "lenguaje de coordenadas" en el cual, por ejemplo, figuren como individuos regiones espacio-temporales finitas, entonces la hipótesis de los cuervos asume una forma tal como "toda región espacio-temporal que contiene un cuervo contiene algo negro"; y aun cuando el número total de cuervos que puedan existir en todos los tiempos sea finito, la clase de las regiones espacio-temporales que contienen un cuervo tiene la potencia del continuo, al igual que la clase de las regiones espacio-temporales que contienen algo no-negro. De este modo, para un lenguaje de coordenadas del tipo considerado, la anterior suposición numérica no tiene sustento. Ahora bien, el uso de un lenguaje de coordenadas puede parecer muy artificial en este ejemplo particular, pero será muy apropiado en otros contextos, por ejemplo, el de las teorías físicas de campo. Por otra parte, la suposición numérica de Hosiasson puede justificarse sobre la base de un "lenguaje de cosa", en el cual los objetos físicos de tamaño finito funcionan como individuos. Por supuesto, aun sobre esta base, para toda hipótesis de la forma "todos los P son Q" queda en pie una cuestión empírica, la de saber si la clase de los no-Q es realmente más numerosa que la clase de los P; y en muchos casos esta cuestión será difícil de resolver.

hipótesis. Esto significa que concebimos la confirmación como una relación semántica[27] entre ciertos objetos[28] extralingüísticos, por una parte, y ciertas oraciones, por la otra. Es posible, sin embargo, concebir la confirmación —de una manera alternativa— como una relación entre dos oraciones, una de las cuales describe los elementos de juicio dados y la otra expresa la hipótesis. Así, en lugar de decir que un obejto *a*, que es un cuervo y es negro (o el hecho de que *a* sea al mismo tiempo un cuervo y negro), confirma la hipótesis de que todos los cuervos son negros, podemos decir que la oración-elemento de juicio "*a* es un cuervo y *a* es negro" confirma la oración-hipótesis (brevemente, la hipótesis) "todos los cuervos son negros". Adoptaremos esta concepción de la confirmación como relación entre oraciones por las razones siguientes: Primera, los elementos de juicio aducidos en apoyo o como crítica de una hipótesis científica siempre se expresan en oraciones que tienen, frecuentemente, el carácter de informes de observaciones; y segunda, resultará muy fructífero proseguir la comparación, mencionada en la sección 2, entre el concepto de confirmación y el de consecuencia lógica. Y así como en la teoría de la relación de consecuencia, es decir, en la lógica deductiva, las premisas de las cuales es una consecuencia una conclusión determinada son concebidas como oraciones, no como "hechos", así también nos proponemos concebir los datos que confirman una hipótesis determinada como dada en forma de oraciones.

La anterior referencia a informes observacionales sugiere una cierta restricción que podría imponerse a las oraciones que expresan los elementos de juicio. En verdad, los elementos de juicio aducidos en apoyo de una hipótesis o teoría científica consisten, en último análisis, en datos accesibles a lo que se llama vagamente observación directa, y tales datos son expresables en la forma de "informes observacionales". En vista de esta consideración, limitaremos las oraciones que expresan los elementos de juicio y que forman el dominio de la relación de confirmación a las oraciones que tengan el carácter de informes observacionales. Para dar un significado preciso al concepto de informe observacional, supondremos que disponemos de un "lenguaje de la ciencia" bien determinado y en función del cual se formularán todas las oraciones en consideración, tanto hipótesis como oraciones que expresen elementos de juicio. Supondremos, además, que este lenguaje contiene, entre otros términos, un "vocabulario observacional" claramente delimitado, formado por términos que designan atributos más o menos directamente observables de cosas y sucesos, tales como, por ejemplo, "negro", "más alto que", "da un color amarillo al quemarse", etc., pero no constructos teóricos tales como "compuesto alifático", "luz polarizada circularmente", "hidrógeno pesado", etcétera.

Entenderemos ahora por *hipótesis* toda oración que pueda expresarse

[27] Para una explicación detallada de este concepto, véase C. W. Morris, *Foundations of the Theory of Signs* (Internat. Encyclopedia of Unified Science, vol. I, Nº 2, Chicago, 1938), y R. Carnap, *Introduction to Semantics* (Cambridge, Mass., 1962), en particular secciones 4 y 37.

[28] En lugar de concebir el primer término de la relación como un objeto o una serie de objetos, podemos concebirlo como un estado de cosas (o quizá como un hecho o una proposición, según Nicod), tal como el estado de cosas que consiste en que *a* sea un cuervo negro, etcétera.

en el lenguaje de la ciencia que hemos supuesto, independientemente de que se trata de una oración generalizada, con cuantificadores, o de una oración particular que se refiera solamente a un número finito de objetos particulares. Un *informe observacional* será concebido como una clase finita (o una conjunción de un número finito) de oraciones observacionales; y una oración observacional como una oración que afirme o niegue que un objeto dado tenga determinada propiedad observable (por ejemplo, "*a* es un cuervo", "*d* no es negro", etc.) o que los objetos de una sucesión dada se encuentren en determinada relación observable (por ejemplo, "*a* está entre *b* y *c*").

Ahora bien, el concepto mismo de observabilidad es relativo, obviamente, a las técnicas de observación utilizadas. Lo inobservable para los sentidos desnudos bien puede ser observable por medio de recursos adecuados, tales como telescopios, microscopios, polariscopios, detectores de mentiras, encuestas Gallup, etc. Si por observación directa entendemos esos procedimientos de observación que no hacen uso de recursos auxiliares, entonces podría decirse que términos de propiedades, tales como "negro", "duro", "líquido", "frío", etcétera, y términos de relaciones, tales como "encima", "entre", "espacialmente coincidentes", etcétera, se refieren directamente a atributos observables: si se concibe la observabilidad en un sentido más amplio, de modo que se admita el uso de determinados instrumentos específicos o de otros recursos, el concepto de atributo observable adquiere mayor extensión. Si en nuestro estudio de la confirmación quisiéramos analizar la manera como las hipótesis y teorías de las ciencias empíricas reciben en última instancia apoyo de "las evidencias de los sentidos", entonces deberíamos exigir que los informes observacionales se refieran exclusivamente a atributos directamente observables. En las partes anteriores de esta sección adoptamos este expediente, para mayor simplicidad y concreción. Pero en realidad, la característica lógica general de la relación que se establece entre una hipótesis y un grupo de enunciados empíricos que la sustentan puede estudiarse independientemente de esta restricción a la observabilidad directa. Todo lo que supondremos aquí es que en el contexto del test científico de una hipótesis o teoría dada, se han convenido ciertas técnicas específicas de observación. Ellas determinan un vocabulario observacional, a saber, un conjunto de términos que designan propiedades y relaciones observables por medio de las técnicas aceptadas. Para nuestros fines es suficiente que estos términos que constituyen el vocabulario observacional estén dados. Se define, pues, una oración observacional simplemente como una oración que afirma o niega que un objeto dado, o una serie de objetos, posee uno de esos atributos observables.[29]

[29] El concepto de oración observacional tiene, en el contexto de nuestro estudio, un carácter y una función lógicos estrechamente afines a los de los conceptos de enunciado protocolar u oración básica, etc., tal como se los usa en muchos estudios recientes del empirismo. Sin embargo, el concepto de oración observacional que se propone en el presente estudio es más liberal en cuanto hace a la discusión de los problemas lógicos del testeo y la confirmación independiente de diversos temas epistemológicos sumamente controvertidos. Así, por ejemplo, no estipulamos que los informes observacionales se refieran a sucesos psíquicos o a percepciones sensoriales (es decir, que deban ser expresados en un vocabulario fenomenológico o de la psico-

Nótese que no exigimos que una oración observacional sea verdadera, ni que se la acepte sobre la base de observaciones efectivas; más bien, una oración de observación expresa algo que es decidible por medio de las técnicas aceptadas de observación. En otras palabras, una oración observacional describe un resultado posible de las técnicas de observación aceptadas; afirma algo que podría ser establecido por medio de esas técnicas. Quizá sería más sugerente la expresión "oración de tipo observacional"; pero por conveniencia hemos preferido la expresión más breve. Una observación análoga, por supuesto, podría hacerse sobre nuestra definición de informe observacional como una clase o una conjunción de oraciones observacionales. Se comprende la necesidad de este modo amplio de concebir las oraciones y los informes observacionales: la confirmación, tal como aquí la entendemos, es una relación lógica entre oraciones, al igual que la consecuencia lógica. Ahora bien, que una oración O_1 sea una consecuencia de otra oración O_2 no depende de que O_1 sea o no verdadera (o de que se sepa si lo es o no); análogamente, los criterios para establecer si un enunciado dado, en la expresión que le dan los términos del vocabulario observacional, confirma una hipótesis determinada, no pueden depender de que los enunciados del informe sean verdaderos, o estén basados en la experiencia real, o de otra condición similar. Nuestra definición de confirmación debe permitirnos indicar qué tipo de elementos de juicio *confirmarían* una hipótesis dada, *si* se dispusiera de ellos; y, evidentemente, de las oraciones que caractericen a tales elementos de juicio sólo puede exigirse que expresen algo que *podría* observarse, pero no necesariamente algo que se haya establecido realmente por la observación.

Puede ser útil llevar un poco más allá la analogía entre la confirmación y la consecuencia lógica. La verdad o falsedad de O_1 es ajena a la cuestión de si O_2 es una consecuencia de O_1 (de si O_2 puede inferirse válidamente de O_1); pero en una inferencia lógica que justifique una oración O_2 mostrando que es una consecuencia lógica de una conjunción de premisas, O_1, podemos estar seguros de la verdad de O_2 sólo si sabemos que O_1 es verdadera. Análogamente, la cuestión de si un informe observacional se encuentra en la relación de confirmación con una hipótesis dada no depende de que el informe enuncie hallazgos observacionales reales o ficticios; pero, para saber si una hipótesis confirmada por cierto informe es adecua-

---

logía introspectiva). Según el concepto de oración observacional adoptado en este estudio, pueden concebirse los "objetos" aludidos en una oración observacional en cualquiera de los sentidos mencionados o de otras maneras. Por ejemplo, podrían ser regiones espacio-temporales u objetos físicos tales como piedras, árboles, etc. (la mayoría de los ejemplos dados en este capítulo son oraciones observacionales pertenecientes a este tipo de "lenguaje de cosa"); todo lo que exigimos es que satisfagan las pocas condiciones muy generales enunciadas antes.

Esas condiciones imponen a las oraciones observacionales y a los informes observacionales ciertas restricciones con respecto a su forma; en particular, ni uno ni otro tipo de oración puede contener cuantificadores. Esta estipulación se desprende de los fines del análisis lógico que aquí realizamos; pero no pretendemos afirmar que esta restricción formal sea indispensable. Por el contrario, es muy posible y quizá deseable admitir también oraciones observacionales que contengan cuantificadores. Introducimos nuestra suposición simplificadora principalmente para evitar considerables complicaciones lógicas en la definición de confirmación.

da o aceptable, es necesario, por supuesto, saber si el informe se basa o no en la experiencia real. Así como una conclusión —de la que se muestra que es verdadera— de una inferencia lógica debe ser (a1) inferida válidamente de (a2) un conjunto de premisas verdaderas, así también una hipótesis, para ser científicamente aceptable, debe ser (b1) confirmada formalmente por (b2) informes dignos de confianza sobre hallazgos observacionales.

El problema central de este ensayo es establecer criterios generales para la relación formal de confirmación mencionada en (b1); el análisis del concepto de informe observacional digno de confianza, que pertenece al campo de la pragmática,[30] cae fuera del ámbito del presente estudio. Pero hay un punto que vale la pena mencionar aquí. Un enunciado que tenga la forma de un informe observacional (por ejemplo, acerca de la posición de la aguja de determinado termógrafo a las 3 de la mañana) puede ser aceptado o rechazado en la ciencia sobre la base de la observación directa, o porque lo confirmen o desconfirmen indirectamente otras oraciones observacionales aceptadas (en el ejemplo, éstas podrían ser oraciones que describieran la curva trazada por la aguja durante la noche); y a causa de esta posibilidad de confirmación indirecta, nuestro estudio también se vincula con la cuestión de la aceptación de hipótesis que tienen, ellas mismas, la forma de informes observacionales.

La concepción de la confirmación como una relación entre oraciones análoga a la relación de consecuencia lógica sugiere otro requisito para la definición tentativa de confirmación. Si bien la consecuencia lógica debe ser concebida como una relación básicamente semántica entre oraciones, algunos lenguajes pueden establecer criterios de consecuencia lógica en términos puramente sintácticos. De modo análogo, puede concebirse la confirmación como una relación semántica entre un informe observacional y una hipótesis; pero el paralelo con la relación de consecuencia sugiere la posibilidad, para ciertos lenguajes, de establecer criterios de confirmación puramente sintácticos. Las consideraciones siguientes desembocarán, en efecto, en una definición de la confirmación basada en el concepto de consecuencia lógica y en otros conceptos puramente sintácticos.

La interpretación de la confirmación como una relación lógica entre oraciones no supone ningún cambio esencial en el problema central del presente estudio. En particular, todas las observaciones hechas en las secciones anteriores pueden ser reformuladas fácilmente de una manera concorde con esta interpretación. Así, por ejemplo, la afirmación de que un objeto a que es un cisne y es blanco confirma la hipótesis "(x) [Cisne$(x) \supset$ Blanco(x)]" puede expresarse diciendo que el informe observacional "Cisne(a) · Blanco(a)" confirma esta hipótesis. De igual modo, la condición de equivalencia puede ser reformulada de la siguiente manera: si un informe observacional confirma cierta oración, entonces también confirma toda oración que sea lógicamente equivalente a esta última. El criterio de Nicod y nuestras razones para rechazarlo pueden ser reformulados siguiendo los mismo lineamiento. Presentamos el concepto de confirmación de Nicod

[30] Se encontrará una explicación del concepto de pragmática en las publicaciones indicadas en la nota 27.

como que alude a una relación entre objetos no-lingüísticos, por una parte, y a oraciones, por la otra, porque este enfoque parecía asemejarse más íntimamente a las propias formulaciones de Nicod[31] y porque nos permitía evitar ciertos tecnicismos que son realmente innecesarios en este contexto.

7. *El criterio de confirmación basado en la predicción. Sus inconvenientes*

Estamos ahora en condiciones de analizar un segundo concepto de la confirmación, que se insinúa en muchas discusiones metodológicas y que tiene muchos visos de plausibilidad. Su idea básica es muy simple: las hipótesis generales, en la ciencia tanto como en el uso cotidiano, tienen el propósito de permitirnos prever sucesos futuros; luego, parece razonable considerar toda predicción apoyada por la observación ulterior como elemento de juicio confirmatorio de la hipótesis sobre la cual se basa, y toda predicción que fracasa, como elemento de juicio desconfirmatorio. Demos un ejemplo. Sea H_1 la hipótesis de que todos los metales se dilatan al calentarse; simbólicamente: "(x) [(Metal (x) · Calentado (x)) $\supset$ Dil (x)]". Si tenemos un informe de observación según el cual un cierto objeto *a* es metálico y se lo calienta, entonces, por medio de H_1, podemos obtener la predicción de que *a* se dilatará. Supongamos que es apoyada por la observación y se describe el resultado en un enunciado de observación adicional. Tendremos entonces el informe de observación total: [Metal (a), Calentado (a), Dil(a)].[32] Este informe sería considerado como un elemento de juicio confirmatorio de H_1, porque su última oración da apoyo a lo que podía predecirse, o inferirse, de las dos primeras oraciones por medio de H_1: más explícitamente, porque puede inferirse la última oración de las dos primeras en conjunción con H_1. Ahora bien, sea H_2 la hipótesis de que todos los cisnes son blancos; simbólicamente: "(x) [Cisne (x) $\supset$ Blanco(x)]"; y consideremos el informe de observación [Cisne(a), $\sim$ Blanco(a)]. Este informe constituiría un elemento de juicio desconfirmatorio de H_2, porque la segunda de sus oraciones contradice (y, por ende, no apoya) a la predicción "Blanco(a)", que puede deducirse de la primera oración en conjunción con H_2; o también, simétricamente, porque la primera oración contradice la consecuencia "$\sim$ Cisne(a)" que puede inferirse de la segunda en conjunción con H_2. Obviamente, cualquiera de estas formulaciones implica que H_2 es incompatible con el informe de observación dado. Estos ejemplos sugieren la siguiente definición general de confirmación:

Criterio de confirmación basado en la predicción: Sea H una hipótesis,

[31] (Agregada en 1964.) En realidad, esto es no correcto; véase la nota 18. Pero, como se ve fácilmente, las objeciones planteadas en este artículo contra el criterio de Nicod mantienen su fuerza también cuando se entiende este criterio en el sentido de que las hipótesis generales son confirmadas o desconfirmadas por proposiciones, no por objetos.

[32] Un informe (observacional), como se recordará, puede representarse por una conjunción o por una clase de oraciones observacionales: en el último caso, lo caracterizamos escribiendo las oraciones entre corchetes; se supone, por conveniencia, que las comillas que se usan normalmente para mencionar las oraciones están absorbidas por los corchetes.

B un informe observacional, es decir, una clase de oraciones observacionales; entonces,

a) se dice que B confirma H si puede dividirse a B en dos subclases mutuamente excluyentes, B_1 y B_2, tales que B_2 no es vacía y toda oración de ésta puede ser deducida lógicamente de B_1 en conjunción con H, pero no de B_1 solamente;

b) se dice que B desconfirma H si H contradice lógicamente a B;[33]

c) se dice que B es neutral con respecto a H si no la confirma ni la desconfirma.[34]

Pero si bien este criterio es muy adecuado como enunciación de condiciones suficientes de la confirmación para las hipótesis del tipo ejemplificado antes, en cambio, es demasiado estrecho para servir como definición general de confirmación. Hablando en términos generales, este criterio cumpliría su propósito si fuera posible concebir todas las hipótesis científicas como que afirman conexiones regulares entre aspectos observables de los fenómenos investigados; por ejemplo, si todas fueran de la forma "siempre que la característica observable P se halla presente en un objeto o una situación, entonces también estará presente la característica observable Q". Pero en realidad, la mayoría de las hipótesis y de las leyes científicas no son de este tipo simple; por lo general, expresan conexiones regulares de características que no son observables en el sentido de la observabilidad directa, ni siquiera en un sentido mucho más liberal. Consideremos, a título de ejemplo, la siguiente hipótesis: "Siempre que la luz polarizada en un plano de longitud de onda λ atraviesa una capa de cuarzo de espesor d, entonces su plano de polarización rota en un ángulo α que es proporcional a d/λ". Supongamos que el vocabulario observacional en el cual deben formularse nuestros informes observacionales contiene de modo exclusivo términos que se refieren a atributos directamente observables. En tal caso, puesto que no es posible decidir por observación directa la cuestión de si determinado rayo de luz está polarizado en un plano y tiene la longitud de onda λ, ningún informe observacional del tipo que admitimos puede suministrar una información de este género. Esto no sería esencial, por sí mismo, si pudiéramos admitir, al menos, que el hecho de que determinado rayo de luz esté polarizado en un plano, etcétera, puede inferirse lógicamente de algún informe observacional posible; pues entonces, a partir de un adecuado informe de este tipo, en conjunción con la hipótesis

[33] Podría parecer más natural estipular que B desconfirma H si puede ser dividido en dos clases mutuamente excluyentes, B_1 y B_2, tales que la negación de al menos una oración de B_2 puede deducirse de B_1 en conjunción con H; pero puede demostrarse que esta condición es equivalente a (*b*).

[34] Las siguientes citas del libro de A. J. Ayer *Language, Truth and Logic* (Londres, 1936) formulan de una manera particularmente clara el concepto de confirmación como predicción exitosa (aunque las dos citas no están explícitamente identificadas por una definición): "...la función de una hipótesis empírica es permitirnos prever la experiencia. Por consiguiente, si una observación a la cual es atinente una proposición dada se ajusta a nuestras expectativas, ... esta proposición está confirmada" (*loc. cit.*, págs. 142-143); "...es característico de una genuina proposición fáctica ...el que algunas proposiciones experienciales puedan ser deducidas de ella en conjunción con ciertas premisas, sin que sea deducible de estas otras premisas solamente". (*loc. cit.*, pág. 26.)

dada, podríamos predecir una rotación del plano de polarización; y a partir de esta predicción, que no está todavía expresada en términos exclusivamente observacionales, cabría esperar que se pudieran inferir otras predicciones en forma de genuinas oraciones observacionales. Pero en realidad, una hipótesis que afirme que un rayo de luz está polarizado en un plano debe ser considerada como una hipótesis general, que implica un número ilimitado de oraciones observacionales; de este modo, no se la puede inferir lógicamente de un adecuado conjunto de hallazgos observacionales, sino que a lo sumo éstos pueden confirmarla. La mejor manera de poner de manifiesto el punto lógicamente esencial es mediante un caso abstracto muy simple. Supongamos que R_1 y R_2 son dos relaciones de un tipo accesible a la observación directa y que el campo de la investigación científica contiene infinitos objetos. Consideremos ahora la hipótesis:

(H) $$(x)[(y)R_1(x,y) \supset (Ez)R_2(x,z)],$$

es decir: siempre que un objeto x está en la relación R_1 con todo objeto y, entonces está en la relación R_2 con, al menos, un objeto z. Esta hipótesis simple tiene la siguiente propiedad: por grande que sea el número de oraciones observacionales dadas, H no nos permite de ella ninguna nueva oración de observación. En verdad, para enunciar la razón de ella en términos sugestivos, aunque no formalmente rigurosos, para hacer una predicción concerniente a un objeto específico a, primero tenemos que saber que a está en la relación R_1 con todo objeto; y esta información necesaria, evidentemente, no puede estar contenida en ningún número finito, por grande que sea, de oraciones observacionales, pues un número finito de éstas a lo sumo puede decirnos que a está en la relación R_1 con un número finito de objetos. Por ende, un informe observacional, que siempre contiene solamente un número finito de oraciones observacionales, nunca puede suministrar una base suficientemente amplia para realizar una predicción por medio de H.[35] Además, aun cuando supiéramos que a está en la relación R_1 con todo objeto, la predicción derivable por medio de H no sería una oración observacional: afirmaría que a está en la relación R_2 con *algún* objeto, sin especificar cuál es y dónde hallarlo. Luego, H es una hipótesis empírica que sólo contiene, además de términos puramente lógicos, expresiones pertecientes al vocabulario observacional, no obstante lo cual las predicciones que la hacen posible no parten de informes observacionales ni conducen a ellos.

Es, por ende, una excesiva simplificación decir que las hipótesis y teorías científicas nos permiten inferir predicciones de experiencias futuras a partir de descripciones de experiencias pasadas. Indiscutiblemente, las hipótesis científicas tienen también la función de permitir predicciones; pero la manera como cumplen esta función, el modo de establecer conexiones lógicas entre informes observacionales, es lógicamente más complejo que una inferencia deductiva. Así, en el último ejemplo, el uso de H para

[35] Para ejemplificar: a podría ser un objeto de hierro que quizás sea un imán; R_1 podría ser la relación de atraer; los objetos en investigación podrían ser objetos de hierro. Entonces, un número finito de informes observacionales según los cuales a atrajo a un trozo de hierro particular es insuficiente para *inferir* que a atraerá a todo trozo de hierro.

realizar predicciones asume la siguiente forma: sobre la base de una serie de ensayos individuales, que muestran que *a* se encuentra en la relación R_1 con tres objetos, *b*, *c* y *d*, podríamos aceptar la hipótesis de que *a* está en la relación R_1 con todos los objetos; o también, en términos de nuestro modo formal de discurso: En vista del informe observacional $[R_1(a, b), R_1(a, c), R_1(a, d)]$, la hipótesis de que $(y) R_1(a, y)$ podría admitirse como confirmada por este informe, aunque no lógicamente inferible de él.[36] Podría llamarse a este proceso una casi-inducción.[37] A partir de la hipótesis así establecida podemos luego derivar, mediante *H*, la predicción de que *a* está en relación R_2 al menos con un objeto. Como señalamos antes, esto no es una oración observacional, y en verdad no puede derivarse de ella ninguna oración observacional, pero puede ser confirmada por una oración observacional adecuada, tal como "$R_2(a, b)$". En otros casos es posible la predicción de reales oraciones de observación; por ejemplo, si la hipótesis dada afirma que $(x)((y) R_1(x, y) \supset (z) R_2(x, z))$, entonces, después de aceptar casi-inductivamente, como antes, que $(y) R_1(a, y)$, podemos obtener, por medio de la hipótesis dada, la oración de que *a* está en la relación R_2 con todo objeto, y de aquí podemos deducir predicciones particulares, tales como "$R_2(a, b)$", que tienen la forma de oraciones observacionales.

Así, la cadena de razonamiento que conduce, de hallazgos observacionales determinados, a la "predicción" de otros nuevos, supone, además de inferencias deductivas, ciertos pasos casi-inductivos, cada uno de los cuales consiste en la aceptación de un enunciado intermedio sobre la base de elementos de juicio confirmatorios, pero no concluyentes —por lo habitual— desde el punto de vista lógico. En la mayoría de las predicciones científicas, este esquema general aparece repetidamente; un análisis del uso predictivo de la hipótesis mencionada antes, concerniente a la luz polarizada en un plano, podría servir como ejemplo. Pero en nuestro contexto, basta con esta explicación general de la estructura de la predicción científica. Ella muestra que una definición general de confirmación que haga referencia a predicciones exitosas es circular; en verdad, para dar suficiente amplitud a la formulación original del criterio de confirmación basado en la predicción, tendríamos que reemplazar la frase "puede deducirse lógicamente" por "puede obtenerse mediante una serie de pasos de deducción y casi-inducción"; y la definición de "casi-inducción", en el sentido anterior, presupone el concepto de confirmación.

Observemos, como resultado marginal de la consideración anterior, que un análisis adecuado de la predicción científica (y, análogamente, de

[36] Así, en el ejemplo dado en la anterior nota al pie, la hipótesis de que el objeto *a* atraerá a todo trozo de hierro podría considerarse como suficientemente bien sustentada por un informe observacional según el cual, en los ensayos, *a* atrajo a los objetos de hierro *b*, *c* y *d*, aunque en modo alguno pueda ser deducida de él.

[37] El prefijo "casi" tiene el propósito de diferenciar este procedimiento de la llamada inducción, de la que se supone habitualmente que es un método para descubrir, o inferir, regularidades generales sobre la base de un número finito de casos. En la casi-inducción, no se "descubre" la hipótesis, sino que debe ser *dada* además del informe observacional. El proceso consiste en la aceptación de la hipótesis, si se la juzga suficientemente confirmada por el informe observacional. Véase también el examen de la sección 2c.

la explicación científica y de la puesta a prueba de hipótesis empíricas) exige el análisis del concepto de confirmación. La razón de esto puede reformularse en términos generales de la siguiente manera: las leyes y teorías científicas, por lo común, vinculan términos que se encuentran en el nivel de las construcciones teóricas abstractas, no en el de la observación directa; y a partir de oraciones observacionales, ninguna inferencia lógica puramente deductiva conduce a enunciados acerca de construcciones teóricas que pueden servir como base para realizar predicciones científicas; los enunciados acerca de construcciones teóricas, por ejemplo, "este trozo de hierro es magnético" o "aquí un rayo de luz polarizada atraviesa un cristal de cuarzo", pueden ser confirmados pero no implicados por informes observacionales. Luego, aunque se base en leyes científicas generales, la predicción de nuevos hallazgos observacionales por medio de otros ya dados es un proceso que supone la confirmación, además de la deducción lógica.[38]

8. *Condiciones de adecuación para toda definición de la confirmación*

Hemos hallado, pues, que son insatisfactorias para una definición general de confirmación las dos concepciones más comunes de ésta, formuladas de manera explícita en el criterio de Nicod y en el de la predicción. Además de este resultado negativo, el análisis anterior también puso de manifiesto ciertas características lógicas de la predicción, de la explicación y del test científicos, y nos llevó al establecimiento de ciertas normas que debe satisfacer una definición adecuada de confirmación. Estas normas incluyen la condición de equivalencia y el requisito de que la definición de confirmación sea aplicable a hipótesis de cualquier grado de complejidad lógica, y no solamente al tipo más simple de condicional universal. Una definición adecuada de confirmación, pues, debe satisfacer varios requisitos lógicos, a cuya consideración pasamos ahora.

Ante todo, se convendrá en que toda oración lógicamente implicada por un informe de observación determinado debe considerarse como confirmada por este informe: la implicación es un caso especial de confirmación. Así, por ejemplo, deseamos poder afirmar que el informe de observación "*a* es negro" confirma la oración (hipótesis) "*a* es negro o gris"; y, para referirnos a uno de los ejemplos dados en la sección anterior, la oración observacional "$R_2(a, b)$" debe ser, ciertamente, un elemento de juicio confirmatorio de la oración "$(Ez)R_2(a, z)$". Llegamos así a la esti-

[38] En el anterior esbozo de la estructura de la predicción científica hemos dejado de lado el hecho de que, prácticamente en todos los casos, cuando se dice que se obtiene una predicción por medio de determinada hipótesis o teoría, se utiliza además un conjunto considerable de teorías auxiliares. Así, la predicción de efectos observables de la deflexión de la luz en el campo gravitacional del Sol sobre la base de la teoría general de la relatividad exige las teorías auxiliares de la mecánica y de la óptica. Pero una consideración explícita de este hecho no afectaría nuestro resultado de que las predicciones científicas, aunque se basen en hipótesis o teorías de forma universal, no tienen un carácter puramente reductivo, sino que contienen también pasos casi-inductivos.

pulación de que toda definición adecuada de confirmación debe asegurar el cumplimiento de:

(8.1) *Condición de implicación.* Toda oración implicada por un informe observacional está confirmada por éste.[39]

Esta condición la sugiere la consideración anterior, pero, por supuesto, ésta no la prueba. Hacer de ella una norma de adecuación de la definición de confirmación significa estipular que una definición propuesta de confirmación será rechazada por lógicamente inadecuada si no está construida de tal manera que (8.1) sea satisfecha incondicionalmente. Una observación análoga se aplica a las siguientes normas adicionales propuestas como criterios de adecuación.

En segundo término, un informe de observación que confirma ciertas hipótesis sería considerado invariablemente como confirmante de toda consecuencia de esas hipótesis. Y en verdad, cualquiera de tales consecuencias no es más que la afirmación del contenido conjunto, total o parcial, de las hipótesis originales y, por ende, se la debe considerar confirmada por todo elemento de juicio que confirme estas últimas. Esto sugiere la siguiente condición de adecuación:

(8.2) *Condición de consecuencia.* Si un informe observacional confirma cada una de las oraciones de una clase *K*, entonces también confirma toda oración que sea una consecuencia lógica de *K*.

Si se satisface (8.2), entonces lo mismo es válido para las dos condiciones especiales siguientes:

(8.2.1) *Condición especial de consecuencia.* Si un informe observacional confirma una hipótesis *H*, entonces también confirma toda consecuencia de *H*.

(8.2.2) *Condición de equivalencia.* Si un informe observacional confirma una hipótesis *H*, entonces también confirma toda hipótesis que sea lógicamente equivalente a *H*.

(8.2.2) se sigue de (8.2.1) en razón de que las hipótesis equivalentes son consecuencias mutuas, una de otra. Así, la satisfacción de la condición de consecuencia implica la de nuestra anterior condición de equivalencia, y ésta pierde su carácter de requisito independiente.

En vista del carácter aparentemente obvio de estas condiciones, es interesante observar que la definición de confirmación en términos de predicciones exitosas, si bien satisface la condición de equivalencia, no cumple con la condición de consecuencia. Consideremos, por ejemplo, la formulación del criterio de predicción dada en la primera parte de la sección anterior. Está claro que si pueden predecirse los hallazgos observacionales B_2 sobre la base de los hallazgos B_1 mediante la hipótesis *H*, la misma predicción puede obtenerse mediante cualquier hipótesis equivalente, pero no, en general, mediante una más débil.

[39] Como consecuencia de esta estipulación, un informe observacional contradictorio —tal como [Negro(*a*), ∼ Negro(*a*)]— confirma cualquier oración, pues tiene como consecuencia cualquier oración. Por supuesto, es posible excluir de manera total los informes observacionales contradictorios mediante una ligera restricción en la definición de "informe observacional". Pero no hay ninguna razón importante para hacerlo.

Por otra parte, toda predicción que pueda obtenerse por medio de *H* también puede establecerse, como es obvio, por medio de cualquier hipótesis más fuerte que *H*, esto es, que implica lógicamente a *H*. Así, mientras que la condición de consecuencia estipula, en efecto, que todo lo que confirma una hipótesis dada también confirma toda hipótesis más débil, la relación confirmación definida en términos de predicción exitosa satisfaría la condición de que todo lo que confirma una hipótesis dada también confirma otra más fuerte.

Pero ¿no es esta "condición conversa de consecuencia", como se la podría llamar, suficientemente razonable, y no se la debería incluir entre nuestras normas de adecuación para la definición de confirmación? La segunda de estas sugerencias puede ser descartada fácilmente: la adopción de la nueva condición, además de (8.1) y (8.2), tendría como consecuencia que todo informe observacional *B* confirmaría cualquier hipótesis *H*. Por ejemplo, si *B* es el informe "*a* es un cuervo" y *H* es la ley de Hooke, entonces, de acuerdo con (8.1), *B* confirma la oración "*a* es un cuervo"; luego *B*, de acuerdo con la condición conversa de consecuencia, confirma la oración más fuerte "*a* es un cuervo y la ley de Hooke es válida"; y finalmente en virtud de (8.2), *B* confirmaría *H*, que es una consecuencia de la última oración. Obviamente, el mismo tipo de argumentación puede aplicarse a todos los otros casos.

Pero, ¿no es cierto, a fin de cuentas, que muy a menudo se considera que los datos que confirman una hipótesis *H* confirman también una hipótesis más fuerte? ¿No es cierto, por ejemplo, que se considera que los hallazgos experimentales que confirman la ley de Galileo o las leyes de Kepler también confirman la ley de la gravitación de Newton? [40] Así es, en efecto; pero ello no justifica la aceptación de la condición conversa de la consecuencia como regla general de la lógica de la confirmación, pues en los casos que acabamos de mencionar la hipótesis más débil está conectada con la más fuerte por un vínculo lógico de un tipo particular: es esencialmente un caso de sustitución de la más fuerte. Por ejemplo, mientras que la ley de la gravitación se refiere a la fuerza existente entre dos cuerpos cualesquiera, la ley de Galileo es una especialización referente al caso en que uno de los cuerpos es la Tierra y el otro un objeto cercano a su superficie. Pero en el caso en que el informe observacional de que *a* es un cuervo confirmaba la ley de Hooke, no se daba esta situación; aquí, la regla de que todo lo que confirma una hipótesis dada confirma también cualquier hipótesis más fuerte, se convierte en un principio totalmente absurdo. Por ende, la condición conversa de la consecuencia no constituye una razonable condición general de adecuación.[41]

[40] Hablando en términos estrictos, es posible deducir la ley de Galileo y las leyes de Kepler sólo si se presuponen ciertas hipótesis adicionales, entre ellas las leyes del movimiento; pero esto no afecta al punto en discusión.

[41] William Barrett, en un artículo titulado "Discussion on Dewey's Logic" (*The Philosophical Review*, vol. 50, 1941, págs. 305 y sigs., en particular pág. 312), plantea algunas cuestiones estrechamente relacionadas con lo que hemos llamado antes la condición de consecuencia y la condición de consecuencia conversa. De hecho alude a esta última (aunque sin mencionarla explícitamente) en un argumento destinado a mostrar que "no toda observación que confirma una oración debe necesariamente

Nos queda por enunciar una tercera condición:[42]

(8.3) *Condición de consistencia.* Todo informe observacional lógicamente consistente es lógicamente compatible con la clase de todas las hipótesis que confirma.

Las dos implicaciones más importantes de este requisito son:

(8.3.1) A menos que un informe observacional sea contradictorio,[43] no confirma ninguna hipótesis con la que no sea lógicamente compatible.

(8.3.2) A menos que un informe observacional sea contradictorio no confirma hipótesis que se contradigan mutuamente.

No habrá dificultad para aceptar el primero de estos corolarios; pero quizá se tenga la impresión de que el segundo —y, por consiguiente, el mismo (8.3)— impone una restricción demasiado severa. Podría señalarse, por ejemplo, que un conjunto finito de mediciones concernientes a los cambios de una magnitud física *x*, asociados con los de otra magnitud *y*, pueden adecuarse —y, por ende, podría decirse que confirman— a varias hipótesis diferentes acerca de la función matemática particular en términos de la cual puede expresar la relación entre *x* e *y*; pero tales hipótesis son incompatibles porque, al menos para un valor de *x*, asignarán a *y* diferentes valores.

Sin duda, es posible liberalizar las normas formales de adecuación en consonancia con estas consideraciones. Esto sería equivalente a abandonar (8.3) y (8.3.2), para conservar solamente (8.3.1). Uno de los resultados de esta medida sería que, cuando un informe observacional lógicamente consistente, *B*, confirma dos hipótesis, no confirmará necesariamente su conjunción, ya que las hipótesis pueden ser incompatibles y, por lo tanto, su conjunción contradictoria; luego, por (8.3.1), *B* no la confirmaría. Esta consecuencia es un tanto extraña intuitivamente, y podríamos sentirnos inclinados a sostener que, si bien (8.3) debe ser abandonada y (8.3.1) conservada, (8.3.2) debe ser reemplazada por el requisito (8.3.3): si una oración observacional confirma dos hipótesis, entonces también confirmará su conjunción. Pero es posible mostrar fácilmente que, en virtud de (8.2) este conjunto de condiciones implica la satisfacción de (8.3.2).

---

confirmar también todas sus consecuencias"; en otras palabras, que no es necesario satisfacer siempre la condición especial de consecuencia (8.2.1). Fundamenta su punto de vista con referencia al "caso más simple: la oración 'C' es una abreviatura de 'A·B', y la observación O confirma 'A', *y por ende* 'C', pero es ajena a 'B', que es una consecuencia de 'C'" (las bastardillas son mías).

Por razones contenidas en el anterior examen de la condición de consecuencia y de la condición de consecuencia conversa, la aplicación de esta última al caso en consideración me parece injustificable, de modo que el ejemplo no prueba la tesis del autor; y en verdad, parece haber todo género de razones para mantener la validez irrestricta de la condición de consecuencia. De hecho, el mismo Barret arguye que "el grado de confirmación de la consecuencia de una oración no puede ser menor que el de la oración misma". Esto es totalmente correcto, pero es difícil ver cómo puede reconciliarse la admisión de este principio con el renunciamiento a la condición especial de consecuencia, que simplemente puede ser considerado como su correlativo para la relación no graduada de confirmación.

[42] Véase una cuarta condición en la nota 47.

[43] Un informe observacional contradictorio confirma toda hipótesis (véase nota 39) y es, por supuesto, incompatible con cada una de las hipótesis que confirma.

Por lo tanto, si la condición (8.3) parece demasiado rigurosa, la alternativa más obvia sería reemplazar (8.3) y sus corolarios por la condición mucho más débil (8.3.1), solamente; y es un problema importante saber si puede construirse una definición intuitivamente adecuada de confirmación que satisfaga a (8.1), (8.2) y (8.3.1), pero no a (8.3). Una de las ventajas de una definición que satisfaga a (8.3) es que pondría límite, por así decirlo, a la fuerza de las hipótesis que pueden ser confirmadas por elementos de juicio dados.[44]

En el resto de este estudio, pues, nos ocuparemos exclusivamente del problema de establecer una definición de confirmación que satisfaga las condiciones formales más severas representadas por (8.1), (8.2) y (8.3) conjuntamente.

La satisfacción de esos requisitos, que pueden ser considerados como leyes generales de la lógica de la confirmación, sólo es, por supuesto, una condición necesaria, no suficiente, para la adecuación de cualquier definición de confirmación que se proponga. Así, por ejemplo, si "*B* confirma *H*" fuera definida como "*B* implica lógicamente *H*", entonces las tres condiciones anteriores quedarán satisfechas, como es obvio; pero tal definición no sería adecuada, porque la confirmación debe ser una relación más amplia que la implicación (a esta última se la podría llamar un caso especial de confirmación *concluyente*). Por consiguiente, una definición de confirmación, para ser aceptable, también debe ser materialmente adecuada: debe presentar una razonable semejanza con la concepción de la confirmación que se halla implícita en el procedimiento científico y en el examen metodológico. Esta concepción es vaga y, hasta cierto punto, bastante confusa, como he tratado de mostrar en partes anteriores de este capítulo; por lo tanto, sería demasiado esperar un pleno acuerdo con respecto a si una definición propuesta de confirmación es materialmente adecuada. Por otra parte, habrá un acuerdo bastante grande sobre ciertos puntos; por ejemplo, ahora probablemente se admita que identificar la confirmación con la implicación, o con el criterio de confirmación de Nicod analizado antes o con cualquier definición de confirmación que haga referencia a una "sensación de evidencia" no constituyen aproximaciones adecuadas a ese concepto de confirmación que interesa a la lógica de la ciencia.

Por otra parte, la corrección del análisis lógico (el cual, en un sentido claro, siempre supone una reconstrucción lógica) de un concepto teórico no puede ser evaluada simplemente por nuestros sentimientos de satisfacción ante determinado análisis que se proponga; y si hay, digamos, dos propuestas alternativas para definir un término sobre la base de un análisis lógico, y si ambas parecen acercarse al significado que se quiere lograr, entonces la elección debe hacerse en gran medida con referencia a aspectos tales como las propiedades lógicas de las dos reconstrucciones, y a la amplitud y la simplicidad de las teorías a las cuales conducen.

[44] Esto me lo señaló el doctor Nelson Goodman. La definición que esbozaremos más adelante, y que satisface las condiciones (8.1) y (8.3), se presta, sin embargo, a ciertas generalizaciones que satisfacen solamente las más liberales condiciones de adecuación consideradas.

9. *El criterio de confirmación basado en la satisfacción*

Como ya señalamos, una definición precisa de la confirmación exige que se haga referencia a un determinado "lenguaje de la ciencia", en el que se supone que están formulados todos los informes observacionales y todas las hipótesis examinadas, y cuya estructura lógica se supone determinada con precisión. Cuanto más complejo sea este lenguaje y más ricos sus medios lógicos de expresión, tanto más difícil será, por lo general, establecer para él una adecuada definición de confirmación. Sin embargo, el problema ha sido resuelto, al menos para ciertos casos: con respecto a lenguajes de una estructura lógica relativamente simple ha sido posible construir una definición explícita de confirmación que satisface todos los anteriores requisitos lógicos y que parece intuitivamente bastante adecuada. Se ha publicado en otra parte[45] una exposición de los detalles técnicos de esta definición; en el presente estudio, que se ocupa de los aspectos lógicos y metodológicos generales del problema de la confirmación, más que de detalles técnicos, trataremos de caracterizar la definición de confirmación obtenida lo más claramente posible con un mínimo de tecnicismos.

Consideremos el caso simple de la hipótesis H: "(x) (Cuervo(x) $\supset$ Negro(x))", donde se supone que "Cuervo" y "Negro" son términos de nuestro vocabulario observacional. Sea B un informe observacional según el cual Cuervo(a) $\cdot$ Negro(a) $\cdot$ $\sim$ Cuervo(c) $\cdot$ Negro(c) $\cdot$ $\sim$ Cuervo(d) $\cdot$ $\sim$ Negro(d). Entonces puede decirse que B confirma H en el siguiente sentido: hay tres objetos mencionados en B, a saber, a, c y d; en lo concerniente a éstos, B nos informa que todos aquellos que son cuervos (es decir, el objeto a, exactamente) también son negros.[46] En otras palabras, de la

[45] En el artículo mío menocionado en la nota 2. La estrategia lógica de los lenguajes a los cuales es aplicable la definición mencionada es la del cálculo funcional inferior, con constantes de individuos y con predicados constantes de cualquier grado. Se supone que todas las oraciones del lenguaje están formadas exclusivamente por constantes de predicados, constantes de individuos, variables de individuos, cuantificadores universales y existenciales para variables de individuos y los conectivos de la negación, la conjunción, la alternación y la implicación. No se permite el uso de variables de predicados ni el signo de identidad.

En cuanto a las constantes de predicados, se supone que todas ellas pertenecen al vocabulario observacional, es decir, que denotan propiedades o relaciones observables por medio de las técnicas aceptadas. (Se supone que los términos de predicados "abstractos" se definen mediante los términos del vocabulario observacional, y por ende, que se los reemplaza por sus *definientia,* de modo que nunca aparecen de manera explícita.)

Como consecuencia de estas estipulaciones, puede caracterizarse un informe observacional simplemente como una conjunción de oraciones del tipo ejemplificado por "$P(a)$", "$\sim P(b)$", "$R(c,d)$", "$\sim R(e,f)$", etc., donde "P", "R", etc., pertenecen al vocabulario observacional, y "a", "b", "c", "d", "e", "f", etc., son los nombres de individuos que denotan objetos específicos. También es posible definir más liberalmente un informe observacional como cualquier oración que no contenga cuantificadores, lo cual significa que, además de conjunciones, se incluyen en los informes observacionales alteraciones y oraciones implicativas formadas a partir del anterior tipo de componentes.

[46] Debo la sugerencia de esta idea al doctor Nelson Goodman; ella dio origen a todas las consideraciones que condujeron finalmente a la definición esbozada más adelante.

información contenida en B podemos inferir que la hipótesis H es verdadera dentro de la clase finita de aquellos objetos mencionados en B.

Apliquemos la misma consideración a una hipótesis de una estructura lógicamente más compleja. Sea H la hipótesis "a todo el mundo le gusta alguien"; en símbolos: "$(x)(Ey)$ Gusta(x,y)", esto es, "para toda (persona) x, existe al menos un y (que no es necesariamente una persona distinta) tal que x gusta de y". (También aquí se supone que "gusta" es un término relacional que aparece en nuestro vocabulario observacional.) Supongamos ahora que se nos presenta un informe observacional B en el cual aparecen los nombres de dos personas, digamos "e" y "f". ¿En qué condiciones diremos que B confirma H? El ejemplo anterior sugiere la respuesta: si de B podemos inferir que H queda satisfecha dentro de la clase finita $[e, f]$, es decir, que dentro de $[e, f]$ todo el mundo gusta de alguien. Esto, a su vez, significa que e gusta de e o de f, y f gusta de e o de f. Así se diría que B confirma H si B implicara el enunciado "e gusta de e o de f, y f gusta de e o de f". Este enunciado será llamado el desarrollo de H para la clase finita $[e, f]$.

El concepto de *desarrollo de una hipótesis H para una clase finita de individuos C* puede ser definido con precisión de manera recursiva; aquí bastará decir que el desarrollo de H para C enuncia lo que H afirmaría si existieran exclusivamente aquellos objetos que son elementos de C. Por ejemplo, el desarrollo de la hipótesis $H_1 =$ "$[(x)(P(x) \vee Q(x))]$" (esto es, "todo objeto tiene la propiedad P o la propiedad Q") para la clase $[a, b]$ es "$[P(a) \vee Q(a)] \cdot [P(b) \vee Q(b)]$" (esto es, "$a$ tiene la propiedad P o la propiedad Q, y b tiene la propiedad P o la propiedad Q"); el desarrollo de la hipótesis existencial H_2: al menos un objeto tiene la propiedad P, es decir, "$(Ex)P(x)$", para la clase $[a, b]$ es "$P(a) \vee P(b)$"; el desarrollo de una hipótesis que no contiene cuantificadores —tal como H_3: "$P(c) \vee K(c)$"— se define como la hipótesis misma, cualquiera que sea la clase de individuos a la que se haga referencia.

Un análisis formal más minucioso basado en consideraciones de este tipo conduce a la introducción de una relación general de confirmación, en dos pasos; el primero consiste en definir una relación especial de confirmación directa según los lineamientos indicados; el segundo, en definir la relación general de confirmación mediante la referencia a la confirmación directa.

Omitiendo detalles de secundaria importancia, podemos resumir las dos definiciones de la siguiente manera:

(9.1 Df.). Un informe observacional B confirma directamente una hipótesis H si B implica el desarrollo de H para la clase de aquellos objetos que se mencionan en B.

(9.2 Df.). Un informe observacional B confirma una hipótesis H si ésta es implicada por una clase de oraciones, cada una de las cuales está confirmada directamente por B.

El criterio expresado en estas definiciones podría recibir el nombre de *criterio de confirmación basada en la satisfacción*, porque su idea básica consiste en considerar que una hipótesis está confirmada por un informe

observacional dado, si se satisface la hipótesis en la clase finita de aquellos individuos que se mencionan en el informe.

Apliquemos ahora las dos definiciones a nuestros últimos ejemplos. El informe observacional B_1: "$P(a) \cdot Q(b)$" confirma directamente (y, por ende, también confirma) la hipótesis H_1 porque implica el desarrollo de H_1 para la clase $[a, b]$ dado antes. La hipótesis H_3 no está confirmada directamente por B, porque su desarrollo, o sea la misma H_3, obviamente no es implicada por B_1. Sin embargo, H_3 es implicada por H_1, la cual es confirmada directamente por B_1; luego, en virtud de (9.2), B_1 confirma H_3. De modo análogo puede verse fácilmente que B_1 confirma directamente H_2.

Por último, nos referiremos al primer ejemplo de esta sección. El informe observacional "Cuervo(a) · Negro(a) · $\sim$ Cuervo(c) · Negro(c) · $\sim$ Cuervo (d) · $\sim$ Negro (d)" confirma (aún directamente) la hipótesis "(x) [Cuervo(x) $\supset$ Negro(x)]", pues implica el desarrollo de esta última para la clase $[a, c, d]$, que puede escribirse de la siguiente manera: "[Cuervo(a) $\supset$ Negro(a)] · [Cuervo(c) $\supset$ Negro(c)] · [Cuervo(d) $\supset$ Negro(d)]".

Ahora es fácil definir la desconfirmación y la neutralidad:

(9.3 Df.). Un informe observacional B desconfirma una hipótesis H si confirma la negación de H.

(9.4 Df.). Un informe observacional B es neutral con respecto a una hipótesis H si B no confirma ni desconfirma H.

En virtud del criterio establecido en (9.2), (9.3) y (9.4), todo informe observacional consistente, B, divide todas las hipótesis posibles en tres clases mutuamente excluyentes: las confirmadas por B, las desconfirmadas por B, y aquellas con respecto a las cuales B es neutral.

Puede mostrarse que la definición de confirmación que aquí proponemos satisface todas las condiciones formales de adecuación contenidas en (8.1), (8.2) y (8.3), así como sus consecuencias. Para la condición (8.2), esto es fácil de ver; para las otras condiciones la prueba es más complicada.[47]

[47] Con respecto a estas pruebas, véase el artículo mencionado en la nota 2. Quiero aprovechar esta oportunidad para señalar y remediar cierto defecto de la definición de confirmación desarrollada en ese artículo y que ha sido esbozada antes. Me percaté de ese defecto en una discusión con el doctor Olaf Helmer.

Se convendrá en que una definición aceptable de confirmación debe satisfacer la siguiente condición adicional, que podría haber sido incluida en las normas lógicas de adecuación establecidas en la sección 8: (8.4). Si B_1 y B_2 son informes observacionales lógicamente equivalentes y B_1 confirma (desconfirma o es neutral con respecto a) una hipótesis H, entonces B_2 también confirma (desconfirma o es neutral con respecto a) H. Esta condición se satisface, en verdad, si se conciben los informes observacionales, según lo hemos hecho en este artículo, como clases o conjunciones de oraciones observacionales. Sin embargo, según indicamos al final de la nota 45, esta restricción de los informes observacionales a una forma conjuntiva no es esencial; en realidad, sólo la hemos adoptado aquí para mayor conveniencia de la exposición, y todos los resultados anteriores, inclusive —en particular— las definiciones y los teoremas de esta sección, siguen siendo aplicables sin cambio alguno, si se definen los informes observacionales como oraciones que no contienen cuantificadores. (En este caso, si "P" y "Q" pertenecen al vocabulario observacional, oraciones tales como "$P(a)$ v $Q(a)$", "$P(a)$ v $\sim Q(b)$", etc., se considerarían como informes obser-

Además, la aplicación de la anterior definición de confirmación no está limitada a las hipótesis que tienen la forma de condicionales universales (como sucede con el criterio de Nicod, por ejemplo), ni a las hipótesis universales en general; se aplica, de hecho, a toda hipótesis que pueda expresarse por medio de términos de propiedad y de relación del vocabulario observacional del lenguaje dado, nombre de individuos, los habituales símbolos de conectivos para "no", "y", "o", "si-entonces", y cualquier número de cuantificadores universales y existenciales.

Por último, según lo sugieren tanto los ejemplos anteriores como las consideraciones generales en las que se basa la adopción de la definición anterior, al parecer hemos obtenido una definición de la confirmación que es también materialmente adecuada, en el sentido de que es una razonable aproximación al significado que suele darse a la confirmación.

Un breve examen de ciertos casos especiales de confirmación puede servir para aclarar este último aspecto de nuestro análisis.

10. *Los conceptos relativo y absoluto de verificación y de refutación*

Si un informe observacional implica una hipótesis *H*, entonces, en virtud de (8.1), confirma *H*. Esto concuerda bastante bien con el concepto habitual de elemento de juicio confirmatorio; de hecho, tenemos aquí un caso extremo de confirmación, el caso en que *B confirma H de modo con-*

---

vacionales.) Esta concepción más amplia de los informes observacionales fue adoptada en el artículo mencionado en la nota 2. Pero resultó que, en este caso, la definición de confirmación resumida antes no satisface en general el requisito (8.4). Así, por ejemplo, los informes observaciones B_1 = "$P(a)$" y B_2 = "$P(a) \cdot [Q(b) \vee \sim Q(b)]$" son lógicamente equivalentes, pero mientras que B_1 confirma (y hasta confirma directamente) la hipótesis H_1 = "$(x)P(x)$", el segundo no la confirma, en esencia, porque no implica "$P(a) \cdot P(b)$", que es el desarrollo de H_1 para la clase de los objetos mencionados en B_2. Puede remediarse esta deficiencia del siguiente modo: El hecho de que B_2 no confirme H_1 obedece obviamente a la circunstancia de que B_2 contiene la constante de individuo "*b*", sin afirmar nada acerca de *b*; el objeto *b* sólo es mencionado en un componente analítico de B_2. Se dirá, pues, que el componente atómico "$Q(b)$" aparece (dos veces) de *manera no esencial* en B_2. Diremos en general que un componente atómico, *A*, de una oración molecular *O* aparece de manera no esencial en *O*, si, en virtud de las reglas del cálculo proposicional, *O*, es equivalente a una oración molecular en la cual no aparece *A* en absoluto. Ahora bien, diremos que un objeto es mencionado de manera no esencial en un informe observacional, si sólo se lo menciona en los componentes de este informe que aparecen en él de manera no esencial. El cálculo proposicional suministra procedimientos mecánicos para decidir si un informe observacional dado menciona un objeto de manera no esencial y para establecer formulaciones equivalentes del mismo informe en las que ningún objeto se mencione de manera no esencial. Por último, digamos que un objeto es mencionado de manera esencial en un informe observacional, si en el mismo se lo menciona pero no sólo de manera no esencial. Ahora reemplazaremos (9.1) por la siguiente definición:

(9.1*a*). Un informe observacional *B* confirma directamente una hipótesis *H* si *B* implica el desarrollo de *H* para la clase de aquellos objetos que se mencionan de manera esencial en *B*.

El concepto de confirmación definido por (9.1*a*) y (9.2) satisface ahora (8.4), además de (8.1), (8.2) y (8.3), aunque los informes observacionales sean concebidos de la manera más amplia caracterizada antes en esta nota.

cluyente; este caso se realiza si y sólo si *B* implica *H*. Diremos entonces, también, que *B verifica H*. De este modo, la verificación es un caso especial de confirmación; es una relación lógica entre oraciones; más específicamente, no es otra cosa que la oración de implicación en un dominio restringido a oraciones observacionales.

Análogamente, diremos que *B desconfirma H de modo excluyente*, o que *B refuta H*, si y sólo si *B* es incompatible con *H*. En este caso, *B* implica la negación de *H* y, por ende, en virtud de (8.1) y (9.3), confirma la negación de *H* y desconfirma *H*. Luego, la refutación es un caso especial de desconfirmación; es la relación lógica de incompatibilidad entre oraciones, en un dominio restringido a las oraciones observacionales.

Es claro que los conceptos de *verificación y refutación*, tales como aquí se los define, son *relativos*. Puede decirse que una hipótesis es verificada o refutada sólo con respecto a determinado informe observacional; y una hipótesis puede ser verificada por un informe observacional pero no por otro. Hay, sin embargo, hipótesis que no pueden ser verificadas, y otras que no pueden ser refutadas por ningún informe observacional, como veremos en seguida. Diremos que una *hipótesis* dada es *verificable* (*refutable*) si es posible construir un informe observacional que verifique (refute) la hipótesis. Que una hipótesis sea verificable o refutable, en este sentido, depende exclusivamente de su forma lógica. Para resumir, pueden distinguirse los casos siguientes:

a) Si una hipótesis no contiene términos cuantificacionales como "todos", "algún" o sus equivalentes simbólicos, entonces es verificable y refutable. Por ejemplo, la hipótesis "el objeto *a* se vuelve azul o verde" está implicada, y por ende verificada, por el informe "el objeto *a* se vuelve azul": y la misma hipótesis es incompatible con el informe "el objeto *a* no se vuelve azul ni verde", que, por ende, la refuta.

b) Una hipótesis puramente existencial (es decir, una hipótesis que puede simbolizarse por una fórmula consistente en uno o más cuantificadores existenciales seguidos de una función oracional [*sentential function*] que no contiene cuantificadores) es verificable, pero no refutable, si —como se supone habitualmente— el universo del discurso contiene un número infinito de objetos. Así, por ejemplo, la hipótesis "hay rosas azules" es verificada por el informe observacional "el objeto *a* es una rosa azul", pero ningún informe observacional finito puede contradecir, y por lo tanto refutar, la hipótesis.

c) Recíprocamente, una hipótesis puramente universal (simbolizada por una fórmula consistente en uno o más cuantificadores universales seguidos de una función oracional que no contiene cuantificadores) es refutable, pero no verificable, para un universo infinito del discurso. Por ejemplo, la hipótesis "(x) [Cisne(x) $\supset$ Blanco(x)]" es refutada completamente por el informe observacional [Cisne(a), $\sim$ Blanco(a)]; pero ningún informe observacional finito puede implicar, y por consiguiente verificar, la hipótesis en cuestión.

d) Las hipótesis que no pueden expresarse por oraciones de uno de los tres tipos mencionados hasta ahora, y que exigen —en este sentido— para su formulación cuantificadores universales y existenciales, por lo general no son verificables ni refutables.[48] Así, la hipótesis "toda sustancia es soluble en algún solvente" —simbólicamente, "$(x)(Ey)$ Soluble (x, y)"— no es implicada por ningún informe observacional ni es incompatible con ningún informe semejante, por grande que sea el número de casos de solubilidad o no solubilidad de sustancias particulares en solventes particulares que pueda registrar el informe. Una observación análoga se aplica a la hipótesis "podéis burlar a algunas personas constantemente", cuya formulación simbólica, "$(Ex)(t)Bl(x,t)$", contiene un cuantificador existencial y otro universal. Pero, por supuesto, todas las hipótesis pertenecientes a esta cuarta clase pueden ser confirmadas o desconfirmadas por informes observacionales adecuados, como vimos en la sección 9 con referencia a la hipótesis "$(x)(Ey)$ Gusta(x, y)".

Hemos presentado esta explicación bastante detallada de la verificación y la refutación no sólo con la esperanza de elucidar aún más el significado de la confirmación y la desconfirmación tales como las definimos antes, sino también con el fin de suministrar una base para establecer una tajante diferenciación entre dos significados de verificación (y análogamente, de refutación) que no siempre han sido distinguidos claramente en exámenes recientes del carácter del conocimiento empírico. Uno de los dos significados de verificación que queremos distinguir aquí es el concepto relativo ya explicado; para mayor claridad, lo llamaremos a veces *verificación relativa.* El otro significado es el que podría recibir el nombre de *verificación absoluta o definitiva.* Este último concepto de verificación no pertenece a la lógica formal, sino más bien a la pragmática: se refiere a la aceptación de hipótesis por observadores. científicos, etc., sobre la base de elementos de juicio apropiados. Hablando en términos generales, podemos distinguir tres fases en el *test* científico de una hipótesis dada (que no se dan necesariamente en el orden en que los registramos aquí). La primera fase consiste en la realización de experimentos u observaciones adecuados, y en la ulterior aceptación de los informes observacionales que enuncian los resultados obtenidos; la fase siguiente consiste en confrontar las hipótesis dadas con los informes observacionales aceptados, esto es, en discernir si éstos constituyen elementos de juicio confirmatorios, desconfirmatorios o no atinentes con respecto a la hipótesis; la fase final consiste en aceptar o rechazar la hipótesis sobre la base de los elementos de juicio confirmatorios o desconfirmatorios constituidos por los informes observacionales aceptados, o en suspender el juicio en espera de nuevos elementos de juicio atinentes al caso.

El presente estudio se ha ocupado casi exclusivamente de la segunda

[48] Un estudio más preciso de las condiciones de la no verificabilidad y no refutabilidad supondría la utilización de tecnicismos que son innecesarios para los propósitos del presente estudio. No todas las hipótesis del tipo descripto en (*d*) no son verificables ni refutables; por ejemplo, la hipótesis "$(x)(Ey)[P(x) \vee Q(y)]$" es verificada por el informe "$Q(a)$", y la hipótesis "$(x)(Ey)[P(x) \cdot Q(y)]$" es refutada por "$\sim P(a)$".

fase. Como hemos visto, esta fase tiene un carácter puramente lógico; las normas de evaluación invocadas —a saber, los criterios de confirmación, desconfirmación y neutralidad— pueden ser formuladas de manera completa en términos de conceptos pertenecientes a la lógica pura.

La primera fase, en cambio, tiene un carácter pragmático; no supone ninguna confrontación lógica de oraciones con otras oraciones. Consiste en realizar ciertos experimentos u observaciones sistemáticos y anotar los resultados. Estos se expresan en oraciones que tienen la forma de informes observacionales, y su aceptación por los científicos se vincula (mediante relaciones causales, no lógicas) con experiencias que se producen en esos tests. Por supuesto, una oración que tenga la forma de un informe observacional puede aceptarse en ciertos casos, no sobre la base de la observación directa, sino porque se halla confirmada por otros informes observacionales establecidos previamente; pero este proceso ejemplifica la segunda fase, que hemos examinado antes. Aquí estamos considerando el caso en que una oración es aceptada directamente, "sobre la base de hallazgos experimentales", no porque se apoye en enunciados establecidos anteriormente.

También la tercera fase puede considerarse de carácter pragmático, es decir, como consistente en una decisión por parte de un científico o un grupo de científicos de aceptar (o rechazar, o dejar en suspenso, según sea el caso) una hipótesis dada, después de determinar la cantidad de elementos de juicio confirmatorios o desconfirmatorios de la hipótesis que contienen la totalidad de las oraciones observacionales aceptadas. Sin embargo, puede intentarse una reconstrucción de esta fase en términos puramente lógicos. Esto exigiría el establecimiento de "reglas de aceptación" generales. Hablando de manera aproximada, estas reglas enunciarían en qué medida una hipótesis dada debe ser confirmada por los informes observacionales aceptados para ser científicamente aceptable ella misma;[49] esto es, las reglas formularían criterios para la aceptación o el rechazo de una hipótesis con referencia al tipo y la cantidad de elementos de juicio confirmatorios o desconfirmatorios contenidos en la totalidad de los informes observacionales aceptados. Quizá tales criterios también hicieran referencia a factores adicionales, tales como la simplicidad de la hipótesis en cuestión, su conformidad con el sistema de las teorías anteriormente aceptadas, etc. Queda en pie por el momento la cuestión acerca de la medida en la cual es posible formular en términos puramente lógicos un sistema satisfactorio de tales reglas.[50]

[49] En un artículo de Felix Kaufmann, "The Logical Rules of Scientific Procedure", *Philosophy and Phenomenological Research,* junio, 1942, se encontrará un estimulante examen de algunos aspectos de lo que hemos llamado reglas de aceptación.

[50] La anterior división del test de una hipótesis empírica en tres fases puede ser útil para la clarificación de la cuestión concerniente a si una concepción empirista de la confirmación implica —y en qué medida— una "teoría de la verdad basada en el concepto de coherencia". Este problema ha sido planteado recientemente por Bertrand Russell, quien en el capítulo X de su *Inquiry into Meaning and Truth* ha hecho una serie de objeciones contra las ideas de Otto Neurath sobre este tema (véanse los artículos mencionados en la nota siguiente), y contra declaraciones hechas por

Sea como fuere, la aceptación de una hipótesis sobre la base de un conjunto suficiente de elementos de juicio confirmatorios será, por lo general, tentativa, y sólo será válida "provisoriamente", es decir, con la reserva de que, si aparecen elementos de juicio nuevos y desfavorables (en otras palabras, si es menester aceptar nuevos informes observacionales que desconfirmen la hipótesis en cuestión), la hipótesis será abandonada.

¿Hay excepciones a esta regla? ¿Hay hipótesis empíricas que puedan establecerse definitivamente, hipótesis de las que podamos estar seguros de que, una vez aceptadas sobre la base de elementos de juicio experienciales, nunca tendrán que ser abandonadas? A las hipótesis de este tipo las llamaremos absoluta o definitivamente verificables; el concepto de refutación absoluta o definitiva será concebido de manera análoga.

Mientras que la existencia de hipótesis relativamente verificables o refutables es un simple hecho lógico, ejemplificado al comienzo de esta sección, la cuestión de la existencia de hipótesis absolutamente verificables o refutables es un problema muy controvertido que ha recibido mucha atención recientemente en los escritos de los empiristas.[51] Puesto que este pro-

---

mí en artículos publicados en *Analysis* en 1935 y 1936. Quisiera agregar aquí unas pocas observaciones, necesariamente breves, sobre este problema.

1) Aunque en los artículos de *Analysis* sostuve, en efecto, que la única interpretación posible de la frase "la oración *O* es verdadera" es "*O* está altamente confirmada por informes observacionales aceptados", rechazo ahora esta opinión. Como ha demostrado la obra de A. Tarski, R. Carnap y otros, es posible definir un concepto semántico de verdad que no sea sinónimo del de confirmación fuerte y que corresponde mucho mejor a lo que habitualmente se llama verdad, sobre todo en lógica, pero también en otros contextos. Por ejemplo, si *O* es una oración empírica cualquiera, entonces *O* o su negación es verdadera en el sentido semántico, pero evidentemente es posible que ni *O* ni su negación estén altamente confirmados por los elementos de juicio disponibles. Afirmar que una hipótesis es verdadera equivale a afirmar la hipótesis misma; por lo tanto, sólo puede establecerse la verdad de una hipótesis empírica en el sentido en que puede establecerse la hipótesis misma: esto es, la hipótesis —y por ende, *ipso facto*, su verdad— puede ser más o menos bien confirmada por elementos de juicio empíricos; no hay otra manera de abordar la cuestión de la verdad de una hipótesis.

A la luz de estas consideraciones, me parece aconsejable reservar el término "verdad" para el concepto semántico; yo diría ahora que las declaraciones de los artículos de *Analysis* tratan de la confirmación. (Para un breve y lúcido examen de las características distintivas de la verdad y de la confirmación, véase R. Carnap, "Wahrheit und Bewährung", *Actes I^er. Congrès Internat. de Philosophie Scientifique*, 135, vol. 4; París, 1936.)

2) Ahora es claro también el sentido de que el test de una hipótesis consiste en confrontar oraciones con oraciones y no con "hechos", o sea que es una cuestión de "coherencia" de la hipótesis y de las oraciones básicas aceptadas: todos los aspectos lógicos del testeo científico, es decir, todos los criterios que gobiernan la segunda y la tercera de las tres fases distinguidas antes, sólo se refieren en verdad a ciertas relaciones entre las hipótesis sometidas a test y otras oraciones (a saber, los informes observacionales aceptados); no se necesita ninguna referencia a "hechos" extralingüísticos. Por otra parte, la primera fase, la aceptación de ciertas oraciones básicas en conexión con ciertos experimentos u observaciones, requiere, por supuesto, procedimientos extralingüísticos; pero esto ha sido explícitamente afirmado por el autor en los artículos ya mencionados. Por lo tanto, es erróneo sostener que las ideas concernientes a la verdad y a la confirmación defendidas por el empirismo lógico contemporáneo suponen una teoría de la verdad basada en la coherencia.

[51] Véase en particular A. Ayer, *The Foundations of Empirical Knowledge* (Nueva York, 1940); véase también el artículo del mismo autor "Verification and

blema sólo se halla vinculado laxamente con el tema de este ensayo, me limitaré aquí a unas pocas observaciones generales.

Supongamos que el lenguaje de la ciencia tiene la estructura general caracterizada y presupuesta en los exámenes anteriores, particularmente en la sección 9. Es razonable esperar, entonces, que sólo las hipótesis relativamente verificables por informes observacionales adecuados puedan ser absolutamente verificables. Por ejemplo, las hipótesis de forma universal, que ni siquiera admiten verificación relativa, no pueden, ciertamente, ser absolutamente verificables. Por grande que sea el número de casos en que una hipótesis semejante haya recibido el apoyo de hallazgos experimentales, siempre es posible que nuevos elementos de juicio la desconfirmen. Por lo tanto, limitemos nuestra búsqueda de hipótesis absolutamente verificables a la clase de aquellas hipótesis que son relativamente verificables.

Supongamos ahora que *H* es una hipótesis de este tipo y que está verificada relativamente, es decir, lógicamente implicada por un informe observacional *B*, y que el mismo es aceptado en la ciencia como descripción del resultado de un experimento u observación. ¿Podemos decir entonces que *H* está verificada absolutamente? Como es evidente, esto depende de que el informe *B* sea aceptado de modo irrevocable o de que sea concebible que pueda sufrir el destino de ser desautorizado más adelante. Así, la cuestión de la existencia de hipótesis absolutamente verificables nos remite a la cuestión de si todos los informes observacionales —o al menos algunos de ellos— se convierten en partes irrevocables del sistema de la ciencia, una vez que se los ha aceptado en lo que respecta a ciertas observaciones o experimentos. No es simplemente una cuestión fáctica; no se la puede responder adecuadamente mediante una descripción de la conducta de los científicos en la investigación. Aquí, como en otros casos de análisis lógico de la ciencia, el problema exige una reconstrucción racional del procedimiento científico, esto es, la construcción de un modelo teórico coherente y amplio de la indagación científica, que luego sirva como sistema de referencia o como patrón para el examen de cualquier investigación científica particular. La construcción del modelo teórico, por supuesto, debe tomar en consideración las características del procedimiento científico real, pero no está determinada por éste, en el sentido en que lo estaría una descripción de algún estudio científico. En verdad, se admite generalmente que los científicos a veces violan las normas del procedimiento científico correcto; además, en pro de la amplitud y la sistematización teóricas, el modelo abstracto deberá contener ciertos elementos idealizados que no pueden estar determinados en detalle por un estudio de cómo trabajan realmente los científicos. Esto es verdad, en particular, de los informes observacionales. Un

Experience", *Proceedings of the Aristotelian Society* de 1937. R. Carnap, "Ueber Protokollsätze", *Erkenntnis*, vol. 3 (1932), y el § 82 de la obra del mismo autor *The Logical Syntax of Language* (Nueva York y Londres, 1937); O. Neurath, "Protokollsätze", *Erkenntnis*, vol. 3 (1932); "Radikaler Physikalismus und 'wirkliche Welt'", *Erkenntnis*, vol. 4 (1934); "Pseudorationalismus der Falsifikation", *Erkenntnis*, vol. 5 (1935); K. Popper, *Logik der Forschung* (véase nota 4); H. Reichenbach, *Experience and Prediction* (Chicago, 1938), cap. III; Bertrand Russell, *An Inquiry into Meaning and Truth* (Nueva York, 1940), en particular los caps. X y XI, y M. Schlick, "Ueber das Fundament der Erkenntnis", *Erkenntnis*, vol. 4 (1934).

estudio de cómo se formulan en la práctica de la investigación científica los informes de laboratorio o las descripciones de otros tipos de hallazgos observacionales presenta interés para la elección de supuestos concernientes a la forma y el carácter de las oraciones observacionales del modelo de un lenguaje de la ciencia; pero evidentemente, tal estudio no puede determinar de manera total la forma que deben tener las oraciones observacionales en el modelo teórico. ni si se las debe considerar irrevocables una vez que se las ha aceptado.

Una analogía quizá contribuya a aclarar esta idea concerniente al carácter del análisis lógico. Supongamos que observamos a dos personas que están jugando un juego en un tablero semejante al del ajedrez, y que no comprendemos su lenguaje; supongamos también que queremos "reconstruir" las reglas del juego. Una mera descripción de lo que hacen los jugadores no bastará para lograrlo; en realidad, ni siquiera debemos necesariamente rechazar una reconstrucción teórica del juego que no caracterice siempre de manera exacta las movidas concretas de los jugadores: debemos admitir la posibilidad de ocasionales violaciones de las reglas. Nuestra reconstrucción estará guiada más bien por el objetivo de lograr un sistema de reglas consistente y amplio que sea lo más simple posible, y al cual se ajuste, al menos en gran medida, la conducta observada en los jugadores. Podemos entonces describir y analizar críticamente toda partida concreta del juego en términos de la norma así obtenida.

El paralelismo es obvio, y también parece claro que en ambos casos la decisión acerca de diversos rasgos del modelo teórico tendrá el carácter de una convención, influida por consideraciones de simplicidad, consistencia y amplitud, y no sólo por el estudio del procedimiento concreto de los científicos en su labor.[52]

Esta observación se aplica en particular a la cuestión que estamos considerando, a saber, si "hay" en la ciencia informes observacionales aceptados irrevocablemente (cuyas consecuencias serían todas, entonces, hipótesis empíricas verificadas absolutamente). La situación se aclara cuando planteamos el problema en esta forma: ¿admitiremos, en nuestra reconstrucción racional de la ciencia, la posibilidad de que ciertos informes observacionales sean considerados como irrevocables, o la aceptación de todos los informes observacionales estará sujeta a la cláusula del "provisoriamente"? Al comparar los méritos de las estipulaciones alternativas, tendríamos que investigar la medida en la cual cada una de ellas es capaz de aclarar la estructura de la indagación científica en términos de una teoría simple y consistente. No nos proponemos entrar a examinar aquí esta cuestión, pero indicaremos que diversas consideraciones dan apoyo a la convención de no aceptar ningún informe observacional como definitivo e irrevocable.[53] Si se elige esta alternativa, entonces ni siquiera las hipótesis implicadas por informes observacionales aceptados son verificadas absolutamente, como

[52] Una clara formulación del sentido en que los resultados del análisis lógico representan convenciones puede encontrarse en los §§ 9-11 y 25-30 de *Logik der Forschung*, de K. Popper.

[53] Véanse en particular las publicaciones de Carnap, Neurath y Popper mencionadas en la nota 51; véase también Reichenbach, *loc. cit.*, sección 9.

tampoco son refutadas absolutamente aquellas hipótesis que son incompatibles con informes observacionales aceptados; en tal caso, en efecto, ninguna hipótesis sería absolutamente verificable o absolutamente refutable. En cambio, si se declaran irrevocables algunas oraciones observacionales —o hasta todas ellas— una vez que se las ha aceptado, entonces las hipótesis implicadas por oraciones observacionales irrevocables serán absolutamente verificadas, y las incompatibles con tales oraciones serán absolutamente refutadas.

Es ahora claro que los conceptos de verificabilidad (o refutabilidad) absoluta y relativa difieren fundamentalmente uno de otro. El no distinguirlos ha sido la causa de considerables malentendidos en discusiones recientes sobre la naturaleza del conocimiento científico. Por ejemplo, la propuesta de Popper de admitir como hipótesis científicas exclusivamente las oraciones que son (relativamente) refutables por informes observacionales adecuados ha sido criticada con argumentos que, en efecto, sustentan la tesis de que no deben concebirse las hipótesis científicas como absolutamente refutables, punto que Popper no ha negado. Sin embargo, según puede verse a través de nuestro anterior examen de la refutabilidad relativa, la propuesta de Popper de limitar las hipótesis científicas a la forma de oraciones (relativamente) refutables supone una restricción muy severa a las formas posibles de hipótesis científicas.[54] En particular, excluye todas las hipótesis puramente existenciales, así como la mayoría de las hipótesis cuya formulación requiere la cuantificación universal y la existencial. Se la puede criticar por esto, ya que en términos de esta reconstrucción teórica de la ciencia parece difícil o totalmente imposible ofrecer una explicación adecuada del carácter y la función de las hipótesis y teorías científicas más complejas.

Lo que dijimos antes acerca de la naturaleza del análisis lógico de la ciencia en general, se aplica al presente análisis de la confirmación en particular: es una propuesta científica de realizar una reconstrucción lógica sistemática y amplia de un concepto básico para la metodología de la ciencia empírica y para la epistemología. La necesidad de una clarificación teórica de este concepto se puso de manifiesto en el hecho de que no se dispone hasta ahora de ninguna explicación teórica general de la confirmación, y de que ciertas concepciones difundidas de la confirmación implican dificultades tan serias que podría dudarse de que pueda lograrse elaborar una teoría satisfactoria de tal concepto.

Hallamos, sin embargo, que el problema puede resolverse: desarrollamos una definición general de confirmación, expresada en términos puramente lógicos, para lenguajes científicos de un carácter específico y relativamente simple. El modelo lógico así obtenido parecía satisfactorio, en el sentido de las normas formales y materiales de adecuación establecidas previamente.

[54] Esto fue señalado por R. Carnap; véase su comentario bibliográfico del libro de Popper en *Erkenntnis*, vol. 5 (1935), y "Testability and Meaning", §§ 25, 26. Una discusión del criterio de refutabilidad de Popper se encontrará por ejemplo, en H. Reichenbach, "Ueber Induktion und Wahrscheinlichkeit", *Erkenntnis*, vol. 5 (1935). O. Neurath, "Pseudorationalismus der Falsifikation", *Erkenntnis*, vol. 5 (1935).

He tratado de exponer los aspectos esenciales del análisis y la reconstrucción de la confirmación propuestos lo más explícitamente posible, con la esperanza de estimular una discusión crítica y de facilitar las ulteriores investigaciones de las diversas cuestiones atinentes a este ámbito de problemas. Entre las cuestiones pendientes que parecen merecer cuidadosa consideración, quisiera mencionar la exploración de conceptos de la confirmación que no satisfagan la condición general de consistencia, la extensión de la definición de confirmación al caso en que se admitan aun oraciones observacionales que contengan cuantificadores, y finalmente la elaboración de una definición de confirmación para lenguajes de una estructura lógica más compleja que la incorporada a nuestro modelo.[55] Los lenguajes de este tipo suministrarían una mayor variedad de medios de expresión y, de este modo, se acercarían más a la gran complejidad lógica del lenguaje de las ciencias empíricas.

POSTSCRIPTUM (1964) SOBRE LA CONFIRMACION

1. *Sobre las paradojas*

Las opiniones expresadas en mi ensayo en lo concerniente a las paradojas aún me parecen correctas: los casos "paradójicos" deben ser considerados como confirmatorios, o positivos; las impresiones en contrario pueden ser atribuibles a factores tales como los sugeridos en la sección 5.2. Varios autores[56] han estado de acuerdo con este juicio, totalmente o en gran medida.

Una serie de comentaristas[57] han argüido, de una manera más o menos

[55] Los lenguajes a los cuales es aplicable nuestra definición tienen la estructura del cálculo funcional inferior, sin signo de identidad. Sería muy deseable ampliar de tal modo la teoría general de la confirmación que se la pudiera aplicar al cálculo funcional inferior con identidad, y aun a cálculos funcionales superiores; pues no parece posible dar una formulación precisa a teorías científicas más complejas sin los medios lógicos de expresión que suministran los cálculos funcionales superiores.

[56] Entre ellos, H. G. Alexander, "The Paradoxes of Confirmation", *The British Journal for the Philosophy of Science*, vol. 9 (1959) 227-233; R. Carnap, *Logical Foundations of Probability* (Chicago, 1950), pág. 469; I. J. Good, "The Paradox of Confirmation", partes I y II, *The British Journal for the Philosophy of Science*, vol. 11 (1960), págs. 145-148; vol. 12 (1961), págs. 63-64; N. Goodman, *Fact, Fiction, and Forecast* (Cambridge, Mass., 1955), págs. 69-73; J. L. Mackie, "The Paradoxes of Confirmation", *The British Journal for the Philosophy of Sciencie*, vol. 13 (1963), págs. 265-277; I. Scheffler, *The Anatomy of Inquiry* (Nueva York, 1963), parte III. En nombre del anti-inductivismo de Popper han planteado críticas, por ejemplo, W. N. Watkins, "Between Analytic and Empirical", *Philosophy*, vol. 32 (1957), págs. 112-131, y "A Rejoinder to Professor Hempel's Reply", *Philosophy*, vol. 33 (1958), págs. 349-355; J. Agassi, "Corroboration versus Induction", *The British Journal for the Philosophy of Science*, vol. 9 (1959), págs. 311-317. Un examen de estas y de otras críticas se hallará en Alexander, *loc. cit.*; Hempel, "A Note on the Paradoxes of Confirmation", *Mind*, vol. 55 (1946), págs. 79-82, y "Empirical Statements and Falsifiability", *Philosophy*, vol. 33 (1958, págs. 342-248; Mackie, *loc. cit.*; Scheffler, *loc. cit.*; R. H. Vincent, "The Paradoxes of Confirmation", *Mind*, vol. 73 (1964), págs. 273-279.

[57] Entre ellos, Alexander, *loc cit.*; Good, *loc. cit.*; D. Pears, "Hypotheticals"

similar a la de Hosiasson-Lindenbaum,[58] que en ciertas suposiciones pueden establecerse diferencias lógicas objetivas entre casos paradójicos y casos no paradójicos de generalizaciones de la forma "todos los *P* son *Q*". La principal suposición requerida apunta al hecho de que hay muchos más no-*Q* que *P* (o, alternativamente, que la probabilidad de que un objeto sea un no-*Q* es mucho mayor que la de ser un *P*). Varios autores presuponen, además, una adecuada teoría de los grados de confirmación o de las probabilidades inductivas, y algunos también suponen que las generalizaciones tienen una probabilidad inicial positiva. En tales suposiciones, se arguye, por ejemplo, que examinar una cosa no-negra para establecer su carencia de la propiedad de ser un cuervo presenta mucho menos riesgo de refutar la generalización "todos los cuervos son negros" que examinar un cuervo para ver si es negro, y que un resultado positivo del primer tipo de test tiene mucho menor importancia o peso que un resultado positivo del segundo (por ejemplo, Pears, quien no aduce una teoría de los grados de confirmación); o que un caso de tipo paradójico aumentará la anterior probabilidad de la generalización en mucho menor medida que un caso no paradójico.

Algunos de estos argumentos me parecen vulnerables a objeciones como las sugeridas en la nota 26 del capítulo I. Pero, y éste es el punto esencial, aunque se hallen establecidas satisfactoriamente, tales diferencias de grado entre los casos paradójicos y los no paradójicos evidentemente no refutan mi conclusión de que los casos paradójicos son confirmatorios. Mi ensayo se ocupa exclusivamente del concepto clasificatorio o cualitativo de confirmación, y no pretende que los diferentes tipos de casos positivos sean todos confirmatorios en el mismo grado o que tengan igual peso al someter a prueba una generalización.

En cuanto a la cuestión pragmática de por qué los casos paradójicos parecen ser no confirmatorios, Pears[59] quizá tenga razón al sugerir que aquellas palabras descriptivas (por ejemplo, "cuervo", "negro", etc.) que usamos normalmente para formular nuestras generalizaciones seleccionan clases que satisfacen (quizá sea mejor decir que comúnmente se cree que satisfacen) la suposición fundamental acerca del tamaño relativo, y que esto a su vez explica, en virtud del tipo de argumentos mencionados, por qué de los casos paradójicos "se piensa que suministran menos confirmación" que los no paradójicos. En verdad, como dice Mackie,[60] hasta podrían explicar por qué para algunas personas el hallazgo de una cosa no negra que no es un cuervo parece carecer de toda importancia como elemento de juicio. Esto bien puede constituir un factor adicional, diferente de los sugeridos en la sección 5.2 de mi artículo, que contribuye en parte a crear la impresión de algo paradójico.[61]

Analysis, vol. 10 (1950), págs. 49-63; G. H. von Wright, *The Logical Problem of Induction* (Oxford, 1957), págs. 122-127.

58 Véase la nota 26.

59 Pears, *loc. cit.*, págs. 51-52. Esto fue sugerido también por la señorita Hosiasson-Lindenbaum en la nota 11 de su artículo.

60 Mackie, *loc. cit.*, págs. 266-267.

61 Cf. también el lúcido examen de estos problemas realizado por Scheffler, *loc. cit.*

2. *Sobre la definición general de confirmación*

Mi definición formal general de confirmación cualitativa me parece ahora demasiado restrictiva. He aquí algunas razones que sustentan este juicio, en orden de importancia creciente:

a) Algunas hipótesis del tipo que abarca mi definición, aunque lógicamente consistentes, no pueden recibir confirmación de ningún informe observacional lógicamente consistente. Por ejemplo, una hipótesis de la forma

$$(x)(\exists y)Sxy \cdot (x)(y)(z)[(Sxy \cdot Syz) \supset Sxz] \cdot (x) \sim Sxx$$

sólo puede ser satisfecha en un dominio infinito; su desarrollo para cualquier clase finita de objetos es contradictorio. En general, ninguna hipótesis científica que implique la existencia de infinitos objetos puede, de acuerdo con mi definición, ser confirmada por un informe observacional. Vale la pena destacar esto, pero no constituye un inconveniente serio de la definición.

b) Mi definición califica de neutrales ciertos tipos de elementos de juicio que normalmente serían considerados como confirmatorios. Así, como ha señalado Canfield,[62] ningún conjunto finito de oraciones del tipo

$$Rab, Rbc, Rcd, Rde, \ldots$$

confirma la hipótesis

$$H_1: (x)(y)Rxy$$

Un informe que no mencione más que los individuos *a* y *b*, por ejemplo, confirma H_1 sólo si implica el desarrollo de H_1 para la clase $\{a, b\}$, es decir, la oración

$$Raa \cdot Rab \cdot Rba \cdot Rbb$$

Y a medida que aumenta el número de los individuos mencionados en un informe observacional, la condición que el informe debe satisfacer para confirmar H_1 es cada vez más restrictiva. Observaciones análogas se aplican al caso de la desconfirmación.

c) Algunos autores[63] han sostenido que la condición de consistencia para la confirmación es demasiado fuerte, por una razón que yo había considerado, aunque luego dejé de lado, en mis comentarios sobre esta condición, en la sección 8: el mismo fenómeno observable puede ser explicado por dos hipótesis compatibles, y el informe observacional que lo

[62] J. Canfield, "On the Paradox of Confirmation", *Metrika*, vol. 5 (1962), págs. 105-118.

[63] En particular Carnap, en su detallada exposición y análisis crítico de mi ensayo, en las secciones 87 y 88 de *Logical Foundations of Probability* (véanse sobre todo las págs. 476-478). Véanse también los comentarios de K. Popper en *The Logic of Scientific Discovery* (Londres, 1959), pág. 374.

describe normalmente sería considerado como confirmatorio de ambas hipótesis. Este punto tiene a mi juicio considerable importancia; pero si es así, entonces es menester abandonar la condición de consecuencia junto con la condición de consistencia. En caso contrario, un informe que confirme dos hipótesis incompatibles confirmaría cualquier consecuencia de las dos y, por ende, cualquier hipótesis.

Por las razones que hemos examinado aquí brevemente, creo que Carnap tiene razón en su juicio de que el concepto de confirmación definido en mi ensayo "no es demasiado amplio, sino evidentemente demasiado estrecho".[64] Por consiguiente, creo que los criterios especificados en mi definición pueden ser suficientes pero no son necesarios para la confirmación de una hipótesis H por un informe observacional B.

Quizás el problema de formular criterios adecuados de confirmación cualitativa pueda abordarse mejor, a fin de cuentas, por medio del concepto cuantitativo de confirmación. Esto lo ha sugerido, en particular, Carnap, quien sostiene que "todo explicatum adecuado para el concepto clasificatorio de confirmación debe estar de acuerdo al menos con un explicatum adecuado para el concepto cuantitativo de confirmación"; esto es, debe haber al menos una función c que sea un explicatum apropiado para el concepto de probabilidad lógica, tal que, cuando B confirma cualitativamente a H, entonces $c(H, B) > c(H, t)$, donde t es el elemento de juicio tautológico o nulo.[65] En otras palabras, en una definición adecuada de probabilidad lógica, la probabilidad de H sobre la base de B debe ser mayor que la probabilidad *a priori* de H cuando B confirma cualitativamente a H.[66] Este principio general lleva también a Carnap a rechazar la condición de consecuencia para la confirmación cualitativa y a restringir la condición de implicación al caso en que H no es una verdad lógica.

Para terminar, examinaré un aspecto muy diferente del problema. De acuerdo con el objetivo enunciado al final de la sección 6, mi definición de confirmación es puramente sintáctica, ya que para los lenguajes formalizados en cuestión el concepto de consecuencia lógica, que aparece en el definiens, es caracterizable en términos puramente sintácticos, como todos los otros conceptos usados en la definición. Pero la confirmación, tanto en su forma cualtitativa como en la cuantiativa, no puede definirse de manera adecuada por medios sintácticos solamente. Esto lo ha aclarado, en particular, Goodman,[67] quien ha mostrado que algunas hipótesis de la forma "$(x)(Px \supset Qx)$" no pueden obtener ninguna confirmación, ni siquiera de oraciones de la forma "$Pa \cdot Qa$". Para aclarar esto adaptaré el ejemplo de Goodman a mi paradigma ornitológico. Representemos "x es un cuervo" por "x es P", y "x es negranco" por "x es Q", donde negranco

[64] Carnap, *loc. cit.*, pág. 479.
[65] Carnap, *loc. cit.*, pág. 472.
[66] Como ha observado Mackie, otros autores conciben la confirmación de acuerdo, más bien, con el "*Principio Inverso*, según el cual una hipótesis h está confirmada por un informe observacional b en relación con el conocimiento básico, si y sólo si el informe observacional se hace más probable agregando la hipótesis al conocimiento básico" (*loc. cit.*, pág. 267; bastardilla del autor).
[67] Goodman, *loc. cit.*, capítulos III y IV.

es todo objeto que haya sido examinado antes de un cierto tiempo *t* y sea negro o no haya sido examinado antes de *t* y sea blanco. Entonces, todo cuervo examinado antes de *t* y negro es un caso formalmente confirmatorio, en el sentido del criterio de Nicod, de la hipótesis "todos los cuervos son negrancos". Pero por grande que sea el número de casos reunidos, éstos no dan ningún apoyo o confirmación a la hipótesis, pues ésta implica que todos los cuervos no examinados antes de *t* —y por consiguiente, en particular, todos aquellos que puedan ser examinados después de *t*— son blancos, y, sin duda, debe considerarse desconfirmada esta consecuencia, no confirmada. Que una hipótesis condicional universal pueda ser confirmada por sus casos positivos, que pueda ser "proyectada" —como dice Goodman— de los casos examinados a los no examinados, dependerá del carácter de sus predicados constituyentes; el uso del predicado "negranco", por ejemplo, excluye tal posibilidad de proyección. Goodman remite la diferencia entre predicados que pueden aparecer en hipótesis proyectables y los que no pueden aparecer de tal modo a su "atrincheramiento", esto es, a la medida en que ellos (o predicados coextensos con ellos) han sido usados en generalizaciones previamente proyectadas; "negranco", por ejemplo, no habiendo sido usado nunca de tal manera, está mucho menos "atrincherado" que términos como "negro", "blanco" y "cuervo". Sobre la base del "atrincheramiento" relativo de los predicados constituyentes, Goodman formula criterios para establecer la proyectabilidad relativa de hipótesis condicionales universales y, por ende, su susceptibilidad a la confirmación por casos formalmente positivos.

Así, la búsqueda de criterios puramente sintácticos de confirmación cualitativa o cuantitativa presupone que las hipótesis en cuestión están formuladas en términos que permiten la proyección, y tales términos no pueden ser individualizados por medios sintácticos solamente. En verdad, la noción de atrincheramiento que usa Goodman para este propósito tiene un carácter claramente pragmático.

Capítulo II

Inconsistencias inductivas [1]

1. *Introducción*

En el estudio filosófico de la inducción, ninguna tarea tiene mayor importancia que la de dar una caracterización clara de los procedimientos inductivos: sólo cuando se haya hecho esto podrá plantearse con sentido el problema de la justificación. Si se concibe la inducción como un tipo peculiar de razonamiento inferencial, su caracterización precisa exigirá, naturalmente, la formulación de reglas claras de inferencia inductiva. En la literatura filosófica, en efecto, se han presentado muchas de tales reglas. Pero algunos tipos muy conocidos de tales reglas, aunque han recibido mucho apoyo aun en escritos recientes sobre el tema, conducen —como puede demostrarse— a inconsistencias lógicas. Esto es tanto más grave cuanto que las reglas defectuosas incluyen algunas que son consideradas como representantes de los tipos fundamentales de correcto razonamiento inductivo. En este capítulo me propongo exponer este defecto de dos tipos familiares de reglas de inducción y examinar las fuentes de las "inconsistencias inductivas" que generan. Las ideas aquí expuestas se basan en gran medida en la obra de otros, y particularmente en la concepción de Carnap de la lógica inductiva y de sus aplicaciones.

2. *Inconsistencias originadas por silogismos estadísticos*

Un tipo de inferencia inductiva que conduce a inconsistencias está representado por el llamado silogismo estadístico y sus variantes.

Un silogismo estadístico[2] es un razonamiento de la forma:

a es *F*

[1] Este capítulo es una reimpresión, con ligeros cambios, del artículo publicado en *Synthese* 12, págs. 439-469 (1960) y que aquí aparece con la amable autorización del Comité Editorial General de esta revista.

Escrito durante mi desempeño, con una beca del gobierno de los Estados Unidos, como Fullbright Research Fellow en la Universidad de Oxford, 1959-1960.

[2] Véase, por ejemplo, D. C. Williams, *The Ground of Induction* (Harvard University Press, 1947) y el examen de la idea en el cap. IV de S. Barker, *Induction and Hypothesis* (Cornell University Press, 1957).

(2.1) La proporción de F que son G es q.
Luego, con probabilidad q, a es G.

En algunas variantes de este modo de razonamiento, la conclusión o también la segunda premisa se expresa en términos no numéricos. Así, por ejemplo, Toulmin[3] presenta como válidos ciertos tipos de razonamientos a los que llama casi-silogismos y que adoptan formas como la siguiente:

a es F

(2.2) La proporción de F que son G es menor que el 2 %.
Luego, casi ciertamente (o probablemente) a no es G.

a es F

(2.3) La proporción de F que son G es ínfima.
Luego, casi ciertamente (o probablemente) a no es G.

Los esquemas de inferencia aquí registrados sólo son aplicables cuando la clase F es finita, pues sólo entonces tiene un significado claro la frase "la proporción de los F que son G". La interpretación frecuencial de la probabilidad estadística, sin embargo, sugiere tipos análogos de razonamiento que no están sujetos a esta restricción. En la teoría matemática actual, las probabilidades estadísticas son concebidas como medidas de conjuntos regidos por ciertos axiomas; una fórmula del tipo "$p(G, F) = r$", que especifica la probabilidad estadística del conjunto G con respecto al conjunto F, afirma, aproximadamente, que la medida de la intersección de G y F, dividida por la medida de F, es igual a r. La aplicación de la teoría matemática a cuestiones empíricas se efectúa por la interpretación frecuencial de la probabilidad estadística, para la cual "$p(G,F) = r$" enuncia la frecuencia relativa a largo plazo con la que un "experimento de azar" de algún tipo especificado, F, realizado por el hombre o por la naturaleza, tiende a dar un resultado del tipo G. Para el caso en que r es cercana a 1, esta interpretación frecuencial es expresada habitualmente en la siguiente forma: Si $p(G,F)$ es muy cercana a 1, entonces si se realiza una sola vez un experimento de tipo F, es prácticamente seguro que se producirá[4] un resultado de tipo G. Puede pensarse que este principio da validez al siguiente esquema de inferencia, en el cual la segunda premisa ya no exige que la clase F se haga finita:

a es F

(2.4) La probabilidad estadística de que un F sea un G es cercana a 1.
Luego, es casi seguro que a es G.

[3] S. Toulmin, *The Uses of Argument* (Cambridge University Press, 1958), págs. 109 y sigs. (Para la forma de conclusión "casi ciertamente, o probablemente, *a* no es *G*", véase pág. 139.)

[4] Esta formulación sigue a la dada por H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, 1946), pág. 150, y por A. Wald, *On the Principles of Statistical Inference* (University of Notre Dame, Indiana, 1942), pág. 2.

Por conveniencia, a todos los diferentes tipos de inferencia mencionados y a otros análogos a ellos los llamaré *silogismos ampliamente estadísticos*, o más brevemente *silogismos estadísticos*. Ahora bien, se ve fácilmente que todos los silogismos ampliamente estadísticos conducen a inconsistencias porque el caso individual *a* que la conclusión asigna a la clase *G* (o al cual la conclusión atribuye la característica o la propiedad *G*) de hecho pertenecerá a diferentes clases F_1, F_2... cuyos miembros serán *G* con diferentes frecuencias relativas o probabilidades estadísticas. Para los razonamientos de la forma (2.1), un ejemplo dado por Barker[5] aclara esto muy bien. Supongamos que Jones es un texano y que el 99 % de los texanos son millonarios; y también que Jones es un filósofo y que sólo el 1 % de los filósofos son millonarios. Entonces, la regla (2.1) permite la construcción de dos silogismos estadísticos, ambos con premisas verdaderas, con las conclusiones incompatibles de que Jones es un millonario con una probabilidad 0,99 y de que lo es con una probabilidad 0,1.

Consideremos el siguiente ejemplo de Toulmin de un casi-silogismo de la forma (2.2):[6]

Petersen es sueco.

(2.5) La proporción de suecos que son católicos es menor que 2 %.

Luego, casi ciertamente, Petersen no es católico.

Supongamos que las premisas de este razonamiento son verdaderas, Entonces, como ha señalado Cooley,[7] las premisas del siguiente casi-silogismo pueden muy bien ser igualmente verdaderas:

Petersen hizo una peregrinación a Lourdes.

(2.6) Menos del 2 % de los que hacen una peregrinación a Lourdes no son católicos.

Luego, casi ciertamente, Petersen es católico.

Así, el esquema de inferencia casi-silogístico puede conducir de premisas verdaderas a conclusiones incompatibles.[8]

Para dar un ejemplo análogo del esquema (2.4), consideremos un conjunto de 10.000 bolillas de las cuales 9.000 son de vidrio y blancas, mientras que las 1.000 restantes son de marfil, una de ellas blanca y las otras 999 negras. Sea *D* un cierto procedimiento para elegir una de las 10.000 bolillas. Supongamos que éste es un procedimiento de azar, de modo que la probabilidad estadística de obtener una bolilla blanca como resultado de *D* será $p(B, D) = 0{,}9001$. Sea el suceso *b* una ejecución par-

[5] Barker, *loc. cit.*, pág. 76.

[6] Toulmin, *loc. cit.*, pág. 109.

[7] J. Cooley, "On Mr. Toulmin's Revolution in Logic", *The Journal of Philosophy*, 56: 297-319 (1959), pág. 305. La formulación del ejemplo de Cooley ha sido modificada ligeramente para adaptarlo mejor al esquema (2.2).

[8] Aunque Toulmin señala repetidamente que los casi-silogismos son *válidos*, luego agrega la observación: "Por supuesto, debe admitirse que sólo pueden hacerse con propiedad casi-silogismos si los datos iniciales a partir de los cuales argumentamos enuncian todo lo que sabemos de importancia para la cuestión que se discute" (*loc. cit.*, pág. 140). Esta observación, la cual implica que el razonamiento (2.5) "sólo puede hacerse con propiedad" si no se sabe que las premisas del casi-silogismo de Cooley son verdaderas, será considerada en la sección 4.

ticular del experimento D. Entonces (2.4) da el siguiente razonamiento con premisas verdaderas:

b es D

(2.7) $p(B, D) = 0{,}9001$
Por consiguiente, es casi seguro que b es B (es decir, que b da una bolilla blanca).

Supongamos ahora que b da una bolilla de marfil. Entonces, b también puede considerarse como un caso de otro experimento, D^*, que consiste en elegir al azar una de las bolillas de marfil del conjunto dado. Pero, en este experimento, la probabilidad de elegir una bolilla no blanca es $p(-B, D^*) = 0{,}999$; y el esquema (2.4) autoriza a realizar el siguiente razonamiento:

b es D^*

(2.8) $p(-B, D^*) = 0{,}999$
Por consiguiente, es casi seguro que b es $-B$ (es decir, que b no da una bolilla blanca).

Nuevamente tenemos un par de razonamientos rivales que se ajustan a la misma regla y parten de premisas verdaderas, no obstante lo cual conducen a conclusiones incompatibles. A pesar de su aparente plausibilidad, pues, la concepción de ciertos tipos de razonamientos estadísticos como si tuvieran la forma de silogismos ampliamente estadísticos es insostenible; pues estos silogismos dan origen a *inconsistencias inductivas* [9] en el siguiente sentido: dado un razonamiento con premisas verdaderas que tiene la forma de un silogismo estadístico, existe en general un razonamiento rival de la misma forma, también con premisas verdaderas, y cuya conclusión es lógicamente incompatible con la del primer razonamiento.

Esto también es verdad de una regla inductiva de un tipo un poco diferente, que figura entre las enumeradas por Black en ensayos que tratan de la justificación de la inducción. Black la formula del siguiente modo:

R: De *La mayoría de los casos de A examinados en una amplia variedad de condiciones han sido B*; *se infiere*: (probablemente) *El próximo A que se encuentre será B*.[10]

Black agrega que los razonamientos inductivos regidos por R varían en "fuerza" según el número y la variedad de los casos favorables aducidos

[9] En un ensayo que trata del uso explicativo y predictivo de los enunciados de probabilidad estadísticos, he llamado a esta peculiaridad de *ambigüedad* de la explicación y la predicción estadísticas; véase "Deductive-Nomological *vs.* Statistical Explanation", en H. Feigl y G. Maxwell (comps.), *Minnesota Studies in the Philosophy of Science*, vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 98-169. Véase también la sección 3.4 del capítulo XII.

[10] M. Black, "Self-Supporting Inductive Arguments", *The Journal of Philosophy*, 55: 718-725 (1958), pág. 720 (bastardillas en el original); véase también del mismo autor "The Inductive Support of Inductive Rules", en M. Black, *Problems of Analysis* (Cornell University Press, 1954), pág. 196.

en la premisa; de modo que "si bien R nos permite *afirmar una cierta conclusión categóricamente...* la fuerza de la afirmación fluctúa según el carácter de los elementos de juicio".[11] En contraste con los silogismos ampliamente estadísticos, pues, la regla *R* conduce a una conclusión que no contiene un calificativo modal como "probablemente" o "ciertamente"; sin embargo, se supone que la conclusión es afirmada con mayor o menor "fuerza". Nuestros anteriores ejemplos muestran que un razonamiento que, de acuerdo con *R*, conduce de premisas verdaderas a una afirmación muy fuerte de una conclusión determinada, generalmente puede ser igualado por un razonamiento rival, gobernado por la misma regla y que a partir de premisas igualmente verdaderas conduce a la afirmación fuerte de la contradictoria de aquella conclusión. En este sentido, la regla *R* genera inconsistencias.

Las formas deductivas de inferencia nunca generan inconsistencias, naturalmente. En particular, para un razonamiento de la forma silogística

a es *F*

(2.9) Todos los *F* son *G*.

a es *G*,

cuyas premisas son verdaderas, no existe ningún razonamiento rival de la misma forma cuyas premisas sean también verdaderas y cuya conclusión sea lógicamente incompatible con la del razonamiento dado: las conclusiones incompatibles sólo pueden ser deducidas de conjuntos incompatibles de premisas, y los conjuntos de premisas verdaderas no son incompatibles.

3. *Probabilidad: ¿calificativo modal o relación?*

Las inconsistencias que acabamos de señalar no demuestran, por supuesto, que sean incorrectos todos los razonamientos basados en información estadística, sino solamente que no es posible considerar como casi-silogísticos tales razonamientos. Tal concepción parece tender a una asimilación formal demasiado estrecha de los razonamientos estadísticos no deductivos con la inferencia deductiva. Así, por ejemplo, puesto que las premisas del silogismo deductivo (2.9) son verdaderas, la conclusión "*a* es *G*" también será "necesariamente" —es decir, como consecuencia lógica— verdadera, y por tanto puede ser afirmada categóricamente. Pero en los razonamientos estadísticos correspondientes, la verdad de las premisas no garantiza la verdad de "*a* es *G*"; y si se insiste, en analogía con el caso deductivo, en formular una oración que la verdad de las premisas nos auto-

[11] Black, "Self-Supporting Inductive Arguments", pág. 720 (bastardillas mías). Black observa que la regla, "tal como está", no es "una regla totalmente aceptable de la inferencia inductiva" (*ibíd.*); pero sostiene que la regla R puede ser usada en un legítimo razonamiento inductivo que dé apoyo a R misma, y parece correcto suponer, pues, que las fallas que encuentra en esta regla no incluyen un defecto tan decisivo como el de generar inconsistencias.

rice a afirmar, puede parecer tentador hacerlo prefijando a "*a* es *G*" una frase calificativa tal como "es prácticamente cierto que", "muy probablemente" o "con probabilidad *r*". Y es esto precisamente lo que se hace cuando se conciben los razonamientos estadísticos como casi-silogísticos.

Se hace claro que esto es erróneo cuando podemos esquematizar el silogismo deductivo (2.9) mediante la misma frase, en la forma:

a es *F*

(3.1) Todos los *F* son *G*.

Luego, ciertamente (o necesariamente) *a* es *G*.

De hecho, esto es lo que hace Toulmin cuando al equivalente silogístico de uno de sus razonamientos casi-silogísticos le da la forma:

Petersen es sueco

(3.2) Ningún sueco es católico.
Luego, ciertamente, Petersen no es católico.[12]

Pero la certeza a la que aquí se alude es, evidentemente, una relación lógica entre las premisas y la conclusión de un razonamiento deductivo: el enunciado "*a* es *G*" es *cierto* o *necesario con respecto a las premisas dadas*, es decir, es lógicamente implicado por ellas. Considerar el término "ciertamente" a la manera de (3.1) y (3.2), como un calificativo aplicable a un solo enunciado, es incorrecto. Si se quiere expresar la fuerza lógica del razonamiento (2.9) con la ayuda del término "cierto" o de otros términos similares, entonces se lo debe hacer en una forma explícitamente relativizada, como la siguiente:

(3.3) "*a* es *G*" es cierto con respecto a (es decir, es lógicamente implicado por) "*a* es *F*" y "todos los *F* son *G*".

Afirmar lo anterior no equivale a negar que la palabra "cierto" y otras similares puedan ser usadas como calificativos de enunciados aislados, en contextos de la forma "es cierto que *p*", "ciertamente *p*", etc. Permítaseme distinguir tres propósitos principales para los que se usan las frases de esta especie: (I) para afirmar que el enunciado particular colocado en lugar de "*p*", o brevemente el enunciado *p*, es una verdad lógico-matemática o quizás una verdad nomológica (es decir, una consecuencia de ciertas leyes de la naturaleza), de modo que estamos autorizados a afirmarla categóricamente y sin reservas; (II) para afirmar que el enunciado *p* es afirmable categóricamente y sin reservas en algún sentido más amplio y evasivo, sentido que se concibe como regido por normas objetivas (algunos harían tal afirmación, por ejemplo, de una clase de presuntas verdades *a priori* concebidas de tal modo que incluyen las verdades de la lógica y de la matemática como una subclase propia); (III) para mostrar, más que declarar, que quien enuncia la frase pretende afirmar el enunciado *p* sin reservas, y quizá con un énfasis especial. Pero si se entiende "ciertamen-

[12] Toulmin, *loc. cit.*, pág. 131.

te" en el primero de estos sentidos, entonces los razonamientos tales como (3.1) y (3.2) son, simplemente, falaces. Lo mismo es válido para el segundo sentido de "ciertamente". Por ejemplo, si se considera que el calificativo se aplica a todas aquellas oraciones que son verdades *a priori*, en algún sentido especificado, y sólo a ellas, entonces bien puede suceder que, en un razonamiento de la forma (3.1), cuyas premisas son verdaderas, la conclusión, aunque también sea verdadera, no sea una verdad *a priori*; luego, en los razonamientos de la forma (3.1), las premisas no garantizan en modo alguno la conclusión. Puede convertirse el esquema (3.1) en una forma correcta de razonamiento si se agrega el prefijo "ciertamente" a ambas premisas; pero el esquema resultante ya no representa el razonamiento silogístico cuya estructura lógica (3.1) pretendía exhibir. Por último, si se toma la palabra "ciertamente" en el tercer sentido, entonces su presencia es tan ajena a la lógica del razonamiento como lo sería la aparición en su lugar de palabras tales como "enfáticamente", "afortunadamente" o "inesperadamente".

En resumen, pues, es simplemente incorrecto representar la fuerza lógica de un razonamiento silogístico en la forma (3.1) o (3.2), donde la palabra "ciertamente" desempeña el papel de un calificativo modal de la conclusión: aquí, ciertamente debe ser concebida como una relación lógica, como en (3.3). El hecho de que la formulación "cierto con respecto a..." que se usa en (3.3) no aparezca en el castellano corriente no es, por supuesto, una falla de la concepción propuesta: precisamente, por haberse aferrado demasiado a las formulaciones usadas en el lenguaje cotidiano, se ha oscurecido la lógica de las inferencias que estamos considerando.

Observaciones análogas pueden aplicarse a los razonamientos estadísticos como los que pretenden aclarar las nociones de silogismo estadístico y de casi-silogismo. En el contexto de tales razonamientos, las frases como "es prácticamente cierto que", etc., así como la expresión de Black "fuerza de aserción" (*strength of assertion*), deben ser concebidas, no como calificativas de la conclusión, sino como representación de una relación lógica existente entre las premisas y la conclusión: ellas indican la medida en la cual las premisas apoyan o confirman la conclusión (no calificada). Así, por ejemplo, los razonamientos cuya estructura pretende exhibir el esquema (2.2) no indican que de las premisas dadas podamos inferir con validez "casi ciertamente *a* no es *G*", sino que esas premisas prestan un fuerte apoyo al enunciado "*a* no es *G*", o que dichas premisas confieren a este enunciado una probabilidad muy alta.

Así, en analogía con (3.3), los razonamientos que (2.2) pretendía representar podrían esquematizarse del siguiente modo:

(3.4) "*a* no es *G*" es casi cierto (o es altamente probable) con respecto a los dos enunciados "*a* es *F*" y "menos del 2 por ciento de los *F* son *G*".

El concepto de probabilidad aquí aducido no es, por supuesto, el estadístico, que representa —como hemos señalado— una relación cuantitativa entre dos tipos o clases de sucesos, *F* y *G*; es lo que Carnap ha llamado

probabilidad lógica o inductiva, o grado de confirmación, concepto que representa una relación lógica entre enunciados. Esta probabilidad inductiva es el concepto fundamental de las teorías de la probabilidad desarrolladas por Keynes, Mazurkiewicz, Jeffreys, von Wright y otros autores. Aún se discute en qué medida el apoyo inductivo que confiere un enunciado-elemento de juicio e a una hipótesis h puede ser representado por un concepto cuantitativo preciso $c(h, e)$ con las características formales de una probabilidad. Sea como fuere, Carnap ha elaborado un riguroso método general para definir tal concepto que es aplicable a lenguajes formalizados con la estructura de un cálculo funcional de primer orden.[13]

Pero el punto fundamental que aquí analizamos es independiente de las perspectivas de elaborar una teoría cuantitativa precisa de la lógica inductiva: para que términos tales como "casi ciertamente", "probablemente" y "con probabilidad r" expresen la fuerza de los razonamientos estadísticos inductivos que hemos considerado, se los debe entender no como calificativos de enunciados aislados, sino como que representan relaciones entre enunciados. Podrían expresarse estas relaciones en la forma (3.4); o también, dentro del marco de una lógica inductiva cuantitativa como la de Carnap, en fórmulas de la forma:

$$(3.5)\quad c(h, e_1, e_2 \ldots e_n) = r$$

las cuales indican que los enunciados ("premisas" inductivas) $e_1, e_2 \ldots, e_n$ confieren conjuntamente la probabilidad lógica r al enunciado h ("conclusión" inductiva).

Para terminar esta breve comparación de la inferencia deductiva con la inductiva, debemos destacar otro punto. Las esquematizaciones (3.3), (3.4) y (3.5) sólo conciernen a las conexiones lógicas entre las premisas y la conclusión, pero en modo alguno a su verdad o falsedad. Ahora bien, puesto que en un razonamiento deductivo la conclusión no puede dejar de ser verdadera si las premisas lo son, pueden usarse las reglas de la inferencia deductiva para efectuar una transición de enunciados dados, de los que se sabe o se piensa que son verdaderos, a otro enunciado que tenga el mismo status; así, con palabras de Carnap,[14] las reglas de la inferencia deductiva permiten adquirir, por decirlo así, nuevos enunciados, sobre la base de los enunciados que ya se poseen. En una inferencia inductiva, en cambio, las "premisas" sólo dan apoyo parcial a la "conclusión", y por ende la verdad no se transfiere automáticamente de las primeras a la segunda. Luego, aunque todas las premisas pertenezcan a la clase de enunciados previamente aceptados o poseídos, la conclusión no puede agregarse a esta clase; sólo puede ser calificada por un número que represente su probabilidad con respecto a las premisas. En lo referente a las "inferencias" o "razonamientos" inductivos, pues, sólo puede hablarse de una "conclusión" *cum grano salis*: no se puede separar la conclusión de las premisas y afirmarla por sí misma cuando las premisas son verdaderas. La cuestión relativa a

[13] Véanse sobre todo sus *Logical Foundations of Probability* (The University of Chicago Press, 1950), sección 100; y su generalización en *The Continuum of Inductive Methods* (The University of Chicago Press, 1952).
[14] Carnap, *Logical Foundations of Probability*, pág. 206.

si la separabilidad de la conclusión de una inferencia deductiva con premisas verdaderas tiene al menos una análoga más débil en el caso de la inferencia inductiva será considerada en la sección final de este ensayo.

La idea de que en el contexto de los razonamientos inductivos la probabilidad debe ser concebida como una relación ha sido extensamente criticada hace poco por Toulmin, quien censura particularmente a Carnap y Kneale por adoptar esta concepción e insiste, por el contrario, en que "probablemente" y otras expresiones afines son calificativos modales de nuestras aserciones. Más específicamente: "Decir 'probablemente *p*' es afirmar cautelosamente y/o con reservas que *p*: *no* es afirmar que estamos dispuestos, tentativamente, a afirmar que *p*." [15] Ahora bien, en el lenguaje corriente, sin duda, la palabra "probablemente" y las afines a ella con frecuencia se usan de esta manera. Podemos distinguir aquí, más precisamente, dos propósitos a los que puede servir el calificativo en frases como "probablemente *p*": (I) Puede mostrar, más que enunciar, en qué medida el que habla desea comprometerse con *p* (si el calificativo tiene la forma "con probabilidad *r*", entonces el cociente $r/(1-r)$ puede indicar las chances en las que, por las razones que fueren, el que habla está dispuesto a apostar a *p*); o bien (II) el calificativo puede indicar la medida en que es racionalmente afirmable o creíble que *p*, donde la afirmabilidad o credibilidad racional se considera regida por normas objetivas. Toulmin no parece optar de manera inequívoca por uno de estos dos significados con que pueden usarse "probablemente" y otras palabras afines. La siguiente declaración, por ejemplo, sugiere el primer significado: "Cuando digo '*O* es probablemente *P*', me comprometo cautelosamente, tentativamente o con reservas con la idea de que *O* es *P* y (de manera igualmente cautelosa) presto mi autoridad a esta idea".[16] Sin embargo, el segundo significado parece aproximarse más a lo que Toulmin tiene *in mente*; así lo sugiere, por ejemplo, su observación: "En realidad, los enunciados referentes a la probabilidad de *p* conciernen, en la práctica, a la medida en que estamos *autorizados* a apostar, admitir, suscribir, jugar nuestra camisa a *p*..." [17]

Pero cuando se los usa en el primer sentido, los calificativos como "probablemente" no pueden servir, claro está, para poner de manifiesto la lógica de un razonamiento estadístico; y si se los entiende en el segundo sentido, entonces se los debe concebir como relativos a los fundamentos dados. En efecto, la credibilidad de una aserción empírica —en total contraste con su verdad o falsedad— depende de los elementos de juicio disponibles: la teoría del flogisto, por ejemplo, era mucho más creíble sobre la base de los elementos de juicio disponibles antes de las investigaciones de Lavoisier que después de ellas. Luego, una frase de la forma "es muy creíble que *p*" (o "probablemente *p*", en el sentido que estamos considerando) no es un enunciado autosuficiente en mayor medida que una frase de la forma "*x* es un número mayor". Con frecuencia las expresiones de la forma "casi ciertamente *p*", "probablemente *p*", etc., tales como se las usa en el lenguaje corriente, pueden ser consideradas como enunciados

[15] Toulmin, *loc. cit.*, págs. 84 y 85 (bastardillas del autor).
[16] Toulmin, *loc. cit.*, pág. 53.
[17] Toulmin, *loc. cit.*, pág. 83 (bastardillas mías).

elípticos que se refieren a los elementos de juicio totales de que se dispone en el momento en que se las pronuncia o en algún otro momento sugerido por contexto. Cuando decimos, por ejemplo, "probablemente no haya vida en la Luna", es de presumir que se hace una referencia tácita a los elementos de juicio disponibles en el presente. Pero si se entiende de este modo el calificativo en la conclusión de un silogismo estadístico, es decir, como relativo a los elementos de juicio totales disponibles en el momento en que se expresa el silogismo, el razonamiento —por supuesto— no es válido: las premisas de (2.2), por ejemplo, no garantizan la conclusión de que, sobre la base de los elementos de juicio *totales* disponibles, es muy probable que *a* no sea *G*. La única concepción razonable que nos queda es la sugerida en (3.4).

Toulmin rechaza enfáticamente esta relativización de la probabilidad e insiste en que la referencia a los elementos de juicio totales sólo es necesaria para estimar o medir la probabilidad de una hipótesis, que en sí misma es una característica no relacional, así como la referencia a elementos de juicio es necesaria para estimar el valor de verdad de una hipótesis.[18] Pero esta analogía es engañosa. Los valores de verdad —la verdad y la falsedad— son características no relacionales de las hipótesis; es decir, una frase de la forma "la hipótesis *h* es verdadera" es un enunciado autosuficiente que no necesita ser completado con la especificación de un conjunto de elementos de juicio. Para estimar si una hipótesis dada, *h*, es verdadera o falsa, tenemos que remitirnos a los elementos de juicio disponibles, llamémoslos *e*, que darán a *h* una confirmación más o menos alta, $c(h, e)$: esta última representa la probabilidad de *h* —o lo que es igual, la probabilidad de que *h* sea verdadera— sobre la base de los elementos de juicio *e*. Esta probabilidad, normalmente, variará según los elementos de juicio, mientras que el valor de verdad de *h* es totalmente independiente de ella. Así, como observamos antes, las frases de la forma "*h* es probable" o "*h* tiene la probabilidad *r*" no son en modo alguno enunciados autosuficientes, y por ende no tiene mayor sentido hablar de estimar la probabilidad de *h*, que el que tendría hablar de estimar si el número 7 es mayor. Y aunque Toulmin dice cosas interesantes acerca del uso corriente de palabras como "probablemente", sus observaciones no dan en modo alguno un significado claro a la noción de probabilidad como concepto no racional.[19] En

[18] Toulmin, *loc. cit.*, págs. 80-81.

[19] Dice lo siguiente (*loc. cit.*, pág. 55): "Sin duda, si digo 'probablemente esté lloviendo' y no está lloviendo, ...yo estaba equivocado"; y luego califica nuevamente de "paradójica e incompatible con nuestras maneras comunes de pensar" la idea de que "si yo digo 'probablemente está lloviendo', el descubrimiento de que no llueve no refuta mi declaración" (*loc. cit.*, pág. 84). Estas observaciones parecen indicar bastante claramente que, en opinión de Toulmin, el enunciado, "no está lloviendo" implica "no se da el caso de que probablemente esté lloviendo". Pero entonces, por contraposición, "probablemente esté lloviendo" implicaría "está lloviendo". Si bien esta concepción daría un fuerte contenido empírico a oraciones de la forma "probablemente *p*", es, por supuesto, en un todo inaceptable. Además, también entra en conflicto con la observación de carácter general que hace Toulmin de que "no se puede especificar ningún suceso que verifique o refute de manera concluyente una predicción a la que sólo se asigna una cierta probabilidad" (*loc. cit.*, pág. 82). De este modo, sus opiniones sobre el contenido y la refutabilidad de los enunciados de probabilidad no relativizados son poco claros.

este caso, el uso corriente ha resultado ser, sin duda, una guía poco confiable.[20]

Tan pronto como, en la esquematización de las inferencias estadísticas, se reconocen los conceptos de probabilidad, casi-certeza, etc., como relacionales y, por consiguiente, se reemplazan los diversos tipos de silogismos ampliamente estadísticos por esquemas de la especie sugerida en (3.4) y (3.5), se disipa un aspecto desconcertante de los razonamientos estadísticos; a saber, la impresión de que los razonamientos estadísticos permiten establecer, sobre la base de premisas verdaderas, pares de conclusiones incompatibles de formas tales como "casi ciertamente (muy probablemente) *a* es *G*" y "casi ciertamente (muy probablemente) *a* no es *G*". Por ejemplo, los dos razonamientos aparentemente conflictuales (2.5) y (2.6) no establecen, como sugiere incorrectamente la concepción de ellos como casi-silogismos, que Petersen casi ciertamente es católico y que casi ciertamente no lo es; más bien, los razonamientos muestran que, con respecto a *un* conjunto de premisas, el enunciado "Petersen es católico" es altamente probable, mientras que su contradictorio es altamente probable con respecto a *otro* conjunto de premisas. Ahora bien, esto no supone una inconsistencia lógica, como no lo supone la observación de que ciertos conjuntos de premisas implican deductivamente el enunciado "Petersen es católico", mientras que otros conjuntos implican deductivamente su contradictorio.

4. *El requisito de los elementos de juicio totales*

Pero si bien la concepción expresada en (3.4) elimina un aspecto desconcertante de los razonamientos estadísticos, no suprime totalmente el problema planteado por las inconsistencias que hallamos en la sección 2. El problema residual no resuelto es el siguiente: si dos conjuntos de enunciados implican deductivamente consecuencias contradictorias, entonces los enunciados de los dos conjuntos no pueden ser todos verdaderos; luego, al menos uno de los razonamientos se basa en algunas premisas falsas. Pero, como hemos señalado, aunque dos conjuntos de enunciados confieran probabilidades muy altas a conclusiones contradictorias, los enunciados de los dos conjuntos pueden ser todos verdaderos. Así, nos enfrentamos con la siguiente cuestión: dados dos razonamientos inductivos válidos cuyas premisas han sido puestas a prueba y aceptadas como presumiblemente verdaderas, pero cuyas conclusiones —atinentes quizás a algún suceso futuro— son lógicamente incompatibles, ¿sobre cuál de ellos debemos basar nuestras expectativas y decisiones? O, con mayor generalidad: sobre la base de diferentes conjuntos de enunciados que consideramos verdaderos puede asignarse a una hipótesis dada, *h* (por ejemplo, una predicción), probabilidades muy diferentes; ¿cuál de ellas debe tomarse como guía para la formación de nuestras creencias concernientes a la verdad de *h* y para tomar decisiones cuyos resultados dependen de que *h* sea o no verdadera?

[20] Se hallará un examen más detallado de las relaciones entre verdad, probabilidad y verificación en Carnap, R., "Truth and Confirmation" y "The Two Concepts of Probability", sección VI; ambos en H. Feigl y W. Sellars (comps.), *Readings in Philosophical Analysis* (Nueva York: Appleton-Century-Crofts, 1949).

Un principio al que hemos aludido antes y que, de hecho, ha sido aceptado tácita o explícitamente por muchos autores de obras sobre el razonamiento inductivo, sugiere la respuesta. Carnap lo llama *el requisito de los elementos de juicio totales* y lo formula del siguiente modo: "En la aplicación de la lógica inductiva a una situación cognoscitiva dada, deben tomarse los elementos de juicio totales como base para determinar el grado de confirmación." [21] Hablando en términos amplios, podemos decir que, según este requisito, el crédito que es racional otorgar a un enunciado en un momento dado debe estar determinado por el grado de confirmación, o de probabilidad lógica, que el enunciado posee sobre la base de los elementos de juicio totales disponibles en ese momento. Alternativamente, puede determinarse tal crédito por su referencia a cualquier parte de los elementos de juicio totales que dé al enunciado el mismo apoyo o probabilidad que éstos; en este caso se dice que la parte omitida de los elementos de juicio totales es *inductivamente ajena* al enunciado, con respecto a los elementos de juicio utilizados.

En lo concerniente a nuestro problema, este principio implica el corolario de que el apoyo que las premisas de un razonamiento estadístico confieren a su conclusión puede servir para determinar el crédito que puede darse racionalmente a dicha conclusión o a las decisiones racionalmente basadas en ella sólo si las premisas constituyen los elementos de juicio totales, *e*, disponibles en ese momento, o una parte de *e* que sustente la conclusión en la misma medida que *e*.

El cumplimiento con el requisito de los elementos de juicio totales resuelve nuestro problema. En efecto: supongamos que estamos ante dos razonamientos estadísticos, uno de los cuales atribuye la casi-certeza a "*a* es *G*" y el otro a "*a* no es *G*". En tal caso, estos razonamientos no pueden satisfacer ambos el requisito de los elementos de juicio totales, pues si así fuera, las probabilidades que sus premisas confieren a "*a* es *G*" y a "*a* no es *G*", respectivamente, serían iguales a las probabilidades que los elementos de juicio totales confieren a esos enunciados; pero el mismo conjunto de elementos de juicio —es decir, los elementos de juicio totales—, si es lógicamente consistente, no puede conferir elevadas probabilidades a dos enunciados contradictorios, pues las dos probabilidades deben sumar 1.

Dicho sea de paso, todo razonamiento *deductivo* cuyas premisas formen parte de los elementos de juicio totales satisfacen trivialmente el requisito de los elementos de juicio totales. En efecto: en este caso las premisas confieren certeza a la conclusión, es decir, le dan la probabilidad lógica 1; pero lo mismo sucede con los elementos de juicio totales disponibles, pues, por hipótesis, incluyen las premisas del razonamiento dado.[22]

Al llegar a este punto, consideremos brevemente una crítica que Ayer[23]

[21] Carnap, *Logical Foundations of Probability*, pág. 221; véase también R. Carnap, "On the Application of Inductive Logic", *Philosophy and Phenomenological Research*, 8: 133-148 (1947-1948), en particular págs. 138-139.

[22] Sobre este punto véase también Carnap, *Logical Foundations of Probability*, pág. 211.

[23] A. J. Ayer, "The conception of probability as a logical relation". S. Körner (comp.), *Observation and Interpretation. Proceedings of the Ninth Symposium of the Colston Research Society* (Nueva York y Londres, 1957), págs. 12-17.

ha dirigido contra el principio de los elementos de juicio totales y, en verdad, contra la concepción de la probabilidad inductiva como una relación lógica entre enunciados. Ayer observa que, según Keynes, Kneale, Carnap y otros autores, los enunciados probabilísticos no son empíricos: si son verdaderos, son necesariamente verdaderos; si son falsos, son necesariamente falsos. Esta característica es particularmente clara en la teoría de Carnap de la probabilidad inductiva, según la cual todo enunciado de la forma "$c(h, e) = r$", que es la forma básica de un enunciado probabilístico inductivo, es analítico o contradictorio. Ahora bien, si tomamos como h una hipótesis determinada, por ejemplo, que nuestro caballo favorito ganará la carrera de mañana, entonces, señala Ayer, podemos asignarle muchas probabilidades diferentes, simplemente tomando en consideración una cantidad cada vez mayor de elementos de juicio atinentes al caso. Pero puesto que cada uno de estos enunciados probabilísticos constituye una verdad necesaria, ninguno de ellos puede ser considerado como superior a los otros. "La adición de más elementos de juicio puede, es cierto, dar una probabilidad mayor o menor al enunciado en el que estamos interesados. Pero... no puede decirse que esta probabilidad sea más o menos correcta que la obtenida mediante los elementos de juicio con que comenzamos".[24] La dificultad señalada se halla estrechamente relacionada con el problema mencionado al comienzo de esta sección; y Ayer destaca la propuesta de Carnap de resolverlo por medio del principio de los elementos de juicios totales. Pero si bien admite que este principio parece estar de acuerdo, en cierta medida, con el sentido común, Ayer pone en duda la posibilidad de justificarlo "sobre la base de los principios de Carnap", precisamente porque un enunciado probabilístico verdadero concerniente a h que no se basa en los elementos de juicio totales, no es menos analítico que otro basado en el requisito de los elementos de juicio totales.

Pero esta exigencia de una justificación del requisito de los elementos de juicio totales en términos de los principios de la lógica inductiva es ajena a la cuestión; pues, como observa Carnap, el principio de los elemente de juicio totales "no es una regla de la lógica inductiva, sino de la metodología de la inducción".[25] Más explícitamente, podemos decir que el principio especifica una condición necesaria, aunque no suficiente, para la racionalidad de las creencias y las decisiones inductivas. También pueden formularse ciertas condiciones de racionalidad para la aplicación del razonamiento deductivo (aunque, como hemos señalado, el requisito de los elementos de juicio totales se satisfacen trivialmente en este caso); por ejemplo, la racionalidad de las creencias exige que si se acepta un conjunto de enunciados como presumiblemente verdadero, o como expresión de creencias presumiblemente verdaderas, entonces debe aceptarse también toda consecuencia lógica de este conjunto. Pero éste no es un principio de la lógica formal. La lógica formal nos dice que si un conjunto dado de enunciados es verdadero, entonces también son verdaderos tales y cuales

[24] Ayer, *loc. cit.*, pág. 14.

[25] Carnap, *Logical Foundations of Probability*, pág. 211. Acerca de la concepción de Carnap de la metodología de la inducción véanse también págs. 202-205 de la misma obra.

enunciados; pero no nos dice cuáles enunciados debemos creer o tomar como base para la acción. En realidad, la noción de aceptar ciertos enunciados, como la noción de elementos de juicio totales, tiene un carácter pragmático y no puede definirse en términos de los conceptos de la lógica formal deductiva o inductiva.

Pero si no puede justificarse el requisito de los elementos de juicio totales por los principios de la lógica formal inductiva, ¿sobre qué base se lo puede propugnar? Bien podría decirse que es simplemente una elucidación parcial de las condiciones que rigen la creencia racional y la elección racional. Así, Carnap construye un ejemplo en el que se viola el requisito y señala correctamente que todo el mundo consideraría esta violación como un serio error en el razonamiento inductivo.[26] Podría agregarse, dentro del mismo espíritu, que si abandonamos este requisito, a veces nos veríamos llevados a dar crédito a enunciados que los elementos de juicio disponibles señalan como falsos. Por ejemplo, podríamos dar crédito a la generalización "de todo huevo incubado nace un pollito" como resultado de limitar nuestros elementos de juicio al subconjunto de los elementos de juicio totales *e* correspondientes solamente a los huevos de gallina, desdeñando la información adicional, también contenida en *e*, acerca de aves que nacen de otros tipos de huevos, la cual mostraría que nuestra generalización es falsa. Y si bien cabe esperar, por supuesto, que razonamientos inductivos a partir de elementos de juicio disponibles a veces nos lleven a dar elevado crédito a enunciados que, sin saberlo nosotros, son de hecho falsos, la racionalidad exige sin duda que no se asigne esta elevada credibilidad a un enunciado que se sabe falso o, más precisamente, a un enunciado que es lógicamente incompatible con elementos de juicio aceptados.

La aplicación práctica del requisito de los elementos de juicio totales presenta considerables dificultades, pues nuestra información total es siempre tan amplia y compleja que no puede ser expresada en dos enunciados con la forma simple de las premisas de esquemas como (3.4). En realidad, es mucho más compleja que el tipo de elementos de juicio considerado en cualquiera de los teoremas de lógica inductiva construidos hasta ahora. Pero, como observa Carnap, un teorema de la lógica inductiva, y cualquier teorema semejante suministra un esquema para realizar razonamientos inductivos válidos, "puede, sin embargo, ser aplicado directamente, siempre que el conocimiento adicional sea, al menos aproximadamente, ajeno a la hipótesis en cuestión".[27] He tratado de mostrar en otra parte[28] que la ciencia empírica, en verdad, nos ofrece diversos razonamientos explicativos y predictivos de un carácter estadístico bastante simple y que satisfacen el requisito de los elementos de juicio totales, al menos en un claro sentido intuitivo.

Como señalamos antes, también Toulmin aduce un principio de los elementos de juicio totales. Si bien insiste en que todos los casi-silogismos, en el sentido que él da al término, son válidos —y de su validez, como de

[26] Carnap, "On the Application of Inductive Logic", pág. 139.
[27] Carnap, *loc. cit.*, pág. 494.
[28] Hempel, "Deductive-Nomological *vs.* Statistical Explanation", sección 11; "Aspects of Scientific Explanation", sección 3.

la de los silogismos deductivos, dice que es "manifiesta" y "no sujeta a duda, por cierto"—,[29] luego observa que "sólo pueden formularse propiamente casi-silogismos si los datos iniciales a partir de los cuales realizamos nuestra argumentación enuncian todo lo que sabemos en lo atinente a la cuestión". "Si no representan más que una parte de nuestro conocimiento atinente al caso, no deberemos razonar categóricamente sino hipotéticamente: 'Dada solamente la información de que Petersen es sueco, podemos concluir que la probabilidad de que sea católico es pequeña'."[30] No está muy claro qué se entiende aquí por validez ni en qué sentido y por cuáles razones un casi-silogismo, aunque válido, "sólo puede formularse propiamente "si satisface el requisito de los elementos de juicio totales. La última parte del pasaje citado parece indicar que, en opinión de Toulmin, la conclusión de un casi-silogismo (incluyendo su calificativo "probablemente", "casi ciertamente", etc.) puede ser afirmada de manera incondicional si se satisface este requisito. Pero entonces estos enunciados probabilísticos presuntamente no relativos parecen ser equivalentes a enunciados probabilísticos relativos enunciados en forma elíptica y referentes a los elementos de juicio totales disponibles; y dentro de esta concepción, sus casi-silogismos serían habitualmente inválidos, como se mostró en la sección 3. Pero como hemos señalado, Toulmin rechaza la interpretación de sus enunciados probabilísticos como elípticos y sostiene, en cambio, que el apoyo que los elementos de juicio totales dan a una hipótesis suministra la mejor *estimación* de *la* probabilidad de la hipótesis. Pero esto deja sin resolver la cuestión de qué es lo que se supone estimado de tal manera; y queda en la oscuridad, como observamos antes, precisamente qué significados atribuye Toulmin a locuciones tales como "h es casi segura", "h es probable", "la probabilidad de h" y "la chance 'real' del cliente de vivir hasta los ochenta".[31]

5. *Inconsistencias generadas por reglas de inducción elementales*

Pasemos ahora a otra clase de presuntas reglas inductivas que generan inconsistencias. Estas reglas presentan especial interés porque se piensa con frecuencia que son las formas más elementales y fundamentales del razonamiento inductivo; las llamaremos, pues, "reglas de inducción elementales".

He aquí dos ejemplos, el primero de los cuales expresa la presunta forma del razonamiento inductivo por enumeración simple:

(5.1) De *todos los casos examinados de A han sido B* inferir *todos los A son B*.[32]

(5.2) Si entre los n casos observados de A, se ha encontrado que m son casos de B, cabe esperar que m/n de los A sean B. Pero mientras

[29] Toulmin, *loc. cit.*, págs. 131-132.
[30] Toulmin, *loc. cit.*, pág. 140.
[31] Toulmin, *loc. cit.*, pág. 71.
[32] M. Black, "The Inductive Support of Inductive Rules", pág. 196.

tanto, continuar buscando otros casos de A y modificar constantemente la razón m/n a medida que se acumulen nuevos datos.[33]

Supongamos ahora que, con el fin de determinar de qué manera varía una cierta magnitud física y (por ejemplo, la longitud de una barra metálica) en función de otra magnitud física x (por ejemplo, la temperatura de la barra), se han medido en n casos los valores asociados de x e y. Sean (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) los pares de valores asociados así establecidos. Los n puntos cuyas coordenadas cartesianas están dadas por estos pares de números pueden estar vinculados con infinitas curvas diferentes C_1, C_2, ..., cada una de las cuales representa los valores de y como una determinada función de los valores de x. Sean $y = F_1(x)$, $y = F_2(x)$,... Ahora bien, sea A la clase de todos los pares de valores asociados físicamente de las magnitudes x e y; entonces, según nuestras suposiciones, es válido afirmar:

(5.3) Todos los n casos examinados de A satisfacen la fórmula "$y = F_1(x)$". Por consiguiente, la regla (5.1) nos prescribe inferir la ley general:

(5.3a) Todos los A satisfacen la fórmula "$y = F_1(x)$".

Pero, según nuestras suposiciones, es igualmente válido afirmar:

(5.4) Todos los n casos examinados de A satisfacen la fórmula "$y = F_2(x)$", lo cual, por (5.1), da la conclusión:

(5.4a) Todos los A satisfacen la fórmula "$y = F_2(x)$"; y así sucesivamente.

De este modo, sobre la base de los mismos datos empíricos, a saber, las n mediciones de valores físicamente asociados x e y, la regla (5.1) da infinitas leyes presuntas diferentes, cada una de las cuales representa a y como una determinada función matemática de x. Además, puesto que no hay dos de las funciones consideradas que sean idénticas, existen ciertos valores de x para los cuales F_1 y F_2, por ejemplo, asignan a y valores diferentes. Luego, las generalizaciones (5.3a) y (5.4a) son lógicamente incompatibles entre sí; y lo mismo otras dos generalizaciones cualesquiera obtenibles por medio de (5.1).

La regla (5.2) produce inconsistencias inductivas del mismo modo. Para darse cuenta de ello basta observar que (5.2) da (5.1) para el caso en que $m = n$; pero puede demostrarse que surgen inconsistencias cuando m es menor que n.

Se aplica esencialmente el mismo razonamiento a la regla básica de inducción de Reichenbach:

(5.5) Dada una parte inicial de n elementos de una sucesión x_i, de la que resulta la frecuencia f^n, y si además no se sabe nada acerca de

[33] M. Black, "'Pragmatic' Justification of Induction", en M. Black, *Problems of Analysis*, pág. 164.

la probabilidad del segundo nivel de la aparición de un cierto límite p, postulamos que la frecuencia $f^i (i>n)$ se acercará a un límite p dentro del margen $f^n \pm \delta$ cuando se prolonga la sucesión.[34]

En efecto, sean los segmentos iniciales los pares (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) obtenidos por medición; entre ellos, la frecuencia relativa de los pares que presentan la relación funcional F_1 es 1; pero es igual la frecuencia relativa de los pares que presentan las relaciones funcionales F_2, F_3, etcétera. Luego, suponiendo que no se sabe nada todavía de lo que Reichenbach llama probabilidades de segundo nivel, la regla nos indica postular que, si se continúa más allá de los n casos iniciales la medición de valores físicamente asociados de x e y, la proporción de pares que satisfacen F_1 se acercará a un límite comprendido dentro de $1 - \delta$; y lo mismo es válido para la proporción de pares que satisfacen F_2, F_3, etcétera. Y aunque no se dé el caso de que cada uno de estos enunciados sobre límites sea lógicamente incompatible con cada uno de los otros, puede verse fácilmente que habrá infinitos pares de enunciados lógicamente incompatibles entre las postulaciones realizadas. Por ende, también la regla (5.5) conduce de premisas verdaderas a un conjunto lógicamente inconsistente de conclusiones.

Las inconsistencias observadas también tienen importancia con respecto a la idea de que todo razonamiento inductivo presupone un principio de uniformidad de la naturaleza que, cuando se lo usa como premisa mayor principal, convierte los razonamientos inductivos en deductivos o "casi-deductivos".[35] Es bien sabido que los intentos para dar una formulación adecuada de este principio presentan serias dificultades. Por ejemplo, es claramente inadecuada la declaración de que lo que ha sucedido en el pasado, en las mismas circunstancias, sucederá también en el futuro. Si se pretende exigir la total identidad de todas las circunstancias concurrentes, entonces la regla es inaplicable, puesto que las mismas circunstancias simplemente no vuelven a darse; si se requiere solamente la igualdad de las circunstancias "importantes", el principio es una perogrullada, pues toda violación aparente del mismo puede atribuirse a una diferencia en algún factor importante aún no reconocido. Una formulación que evita estos inconvenientes y que también parecería expresar mucho más precisamente la intención del principio de uniformidad es la siguiente:

(5.6) Una generalización apoyada por todos los casos examinados hasta ahora también será apoyada por todos los casos futuros.

[34] H. Reichenbach, *The Theory of Probability* (University of California Press, 1944), pág. 446.

[35] Esta idea, que nos es familiar por la obra de Mill, ha sido propugnada recientemente, por ejemplo, por H. G. Alexander en su contribución al simposio "Convention, Falsification and Induction" en *The Aristotelian Society*, Supplementary Volume 34 (Londres, 1960). Alexander destaca, sin embargo, que varias de tales presuposiciones están implícitas en el razonamiento inductivo, y sugiere que, para tomarlas en cuenta explícitamente, el razonamiento inductivo en la ciencia tendría que adoptar una "forma casi deductiva: 'Casi deductiva' porque es imposible enunciar estas presuposiciones en una forma completamente precisa" (*loc. cit.*, pág. 140).

Pero este principio es contradictorio, pues cuando se lo aplica a nuestro ejemplo, implica que todos los pares de valores físicamente asociados de x e y satisfacen la fórmula "$y = F_1(x)$"; pero también satisfacen la fórmula "$y = F_2(x)$", etcétera, ya que los n pares medidos hasta ahora satisfacen todas esas fórmulas.

El método que hemos usado para engendrar inconsistencias por medio de reglas elementales de inducción es afín al empleado por Goodman al plantear su "nuevo enigma de la inducción".[36] Uno de los ejemplos característicos de su enfoque es el siguiente: Supongamos que, de acuerdo con todos los elementos de juicio disponibles en un cierto tiempo t, todas las esmeraldas examinadas hasta ahora (esto es, antes de t) son verdes. Luego, según las concepciones corrientes de la confirmación, los elementos de juicio totales apoyan la generalización h_1: "Todas las esmeraldas son verdes". Ahora bien, llamemos "verdul" a un predicado que se aplica a los objetos examinados antes de t si son verdes, y a otros objetos si son azules. Luego, de acuerdo con los elementos de juicio totales disponibles en t, todas las esmeraldas observadas hasta ahora son verdules. Luego, los elementos de juicio totales también dan sustento a la generalización h_2: "Todas las esmeraldas son verdules". Pero cuando se aplican a esmeraldas examinadas después de t, las dos hipótesis brindan las predicciones conflictuales de que todas las esmeraldas serán verdes y de que todas serán verdules, y por ende, azules. Goodman observa: "Así, aunque sabemos bien cuál de las dos predicciones incompatibles recibe genuina confirmación, según nuestra presente definición ambas se hallan igualmente bien confirmadas".[37] Sugiere que los elementos de juicio totales confirman h_1 y no h_2 porque la primera es un "enunciado legal" (esto es, tiene las características de una ley, excepto que puede ser falsa), mientras que la segunda no; y porque sólo un enunciado legal puede recibir confirmación de sus casos establecidos. De este modo, surge el nuevo problema de la inducción, a saber, el de enunciar claramente "qué distingue a las hipótesis legales o confirmables de las accidentales o no confirmables".[38] Goodman observa que sólo en la medida en que pueda resolverse este problema podremos establecer una distinción entre inferencias inductivas válidas y no válidas; y luego esboza su "teoría de la proyección", que establece una distinción entre hipótesis confirmables y no confirmables en términos del "refuerzo" de los predicados utilizados en su formulación.[39]

Aunque Goodman realiza su examen en términos de reglas de confirmación, se ve fácilmente que sus pares de hipótesis también pueden utilizarse para demostrar que las reglas de inducción elementales mencionadas antes pueden conducir de un conjunto consistente de elementos de juicio a un conjunto inconsistente de conclusiones: aquí yace la afinidad del argumento de Goodman con el que usamos al comienzo de esta sección. El último, sin embargo, parece agregar una nueva faceta al importante

[36] Véase N. Goodman, *Fact, Fiction, and Forecast* (Harvard University Press, 1955), págs. 73 y sigs.
[37] Goodman, *loc. cit.*, pág. 75.
[38] Goodman, *loc. cit.*, pág. 80.
[39] Goodman, *loc. cit.*, cap. IV.

problema planteado por Goodman. En efecto, podríamos sentirnos inclinados a admitir que una generalización como "todas las esmeraldas son verdules" no es legal y que su aplicabilidad a casos todavía no examinados no está garantizada por los casos previamente establecidos; pero entre las generalizaciones conflictuales que pueden obtenerse a la manera de nuestro anterior ejemplo, hay muchas que parecerían ser igualmente legales y, por ende, igualmente capaces de recibir confirmación de sus casos; y si esto es así, entonces ninguna de estas generalizaciones incompatibles quedaría descartada por restringir las conclusiones inductivas permisibles a enunciados legales. (En verdad, las reglas de Goodman no pretenden establecer una diferencia entre hipótesis bien confirmadas pero legales e incompatibles.)

Supongamos, por ejemplo, que los pares de valores asociados de x e y medidos hasta ahora son: $(0, -1)$; $(1, 0)$; $(2, 1)$. Estos satisfacen las siguientes generalizaciones, entre otras:

$$(5.7)\quad y = (x-1);\ y = (x-1)^3;\ y = (x-1)^5;\ldots$$

$$y = \cos \pi \left(1 - \frac{x}{2}\right);\ y = (x-1)^2 \cos \pi \left(1 - \frac{x}{2}\right);\ldots$$

$$y = (x-1)^4 \cos \pi \left(1 - \frac{x}{2}\right);\ldots$$

Cada una de estas generalizaciones incompatibles dos a dos representa, creo, un enunciado absolutamente legal, capaz de confirmación por casos establecidos. Luego, restringiendo el uso de nuestras reglas de inducción elementales a los casos en que la conclusión es un enunciado legal, podemos eliminar las inconsistencias inductivas del tipo de las construidas por Goodman, pero aún nos quedarán conjuntos inconsistentes de hipótesis del tipo ejemplificado por (5.7).

En las discusiones filosóficas acerca de la justificación de los procedimientos inductivos, las reglas del tipo considerado en esta sección a menudo son vistas como formulaciones esencialmente adecuadas, aunque quizás un poco simplificadas, de normas de razonamiento inductivo;[40] veremos ahora que el problema de la justificación ni siquiera se plantea para estas reglas de inducción elementales, pues conducen a inconsistencias lógicas, de modo que violan el requisito, indudablemente mínimo, que toda regla de procedimiento científico debe llenar, antes de que se plantee el problema de su justificación.

[40] Black, por ejemplo, formula sus diversas reglas de inducción para suministrar una caracterización clara de los principios o políticas cuya justificación está en duda; y el ingenioso argumento de Reichenbach tendiente a dar una justificación de la inducción trata específicamente de su regla, que mencionamos antes, y a la que considera como el principio fundamental del procedimiento inductivo.

¿Las inconsistencias que encontramos aquí son atribuibles nuevamente a una violación del requisito de los elementos de juicio totales? A primera vista esto parece poco plausible, pues las reglas (5.1) y (5.2), así como el principio (5.6), incluyen lo que parece ser una versión simple de este requisito, a saber, la condición de que la información dada abarque todos los casos examinados hasta ahora; y puede entenderse que la regla (5.5) presupone también esta condición. En los ejemplos examinados de contradicciones generadas por reglas de inducción elementales también se suponía satisfecha la condición aludida.

Sin embargo, como lo muestran las paradojas de la confirmación,[41] el concepto de "casos" de una hipótesis no singular no es en modo alguno claro, como podría parecer, y hay buenas razones para pensar, pues, que el requisito de los elementos de juicio totales no puede ser expresado adecuadamente por la condición de que los elementos de juicio deben incluir todos los casos observados hasta el momento. Y, en verdad, con referencia a otros dos ejemplos construidos por Goodman, Carnap ha sostenido que ellos contienen una violación del requisito de los elementos de juicio totales.[42] Para el ejemplo mencionado antes, la objeción de Carnap adoptaría esta forma: en el caso de la predicción de que la próxima esmeralda será verdul, se sabe más que el hecho de que las esmeraldas observadas hasta ahora son todas verdules, es decir, que fueron examinadas antes de *t* y eran verdes o no fueron examinadas antes de *t* y eran azules; se sabe que todas fueron examinadas antes de *t*. Y al no incluir esta información en los elementos de juicio, se viola el requisito de los elementos de juicio totales.

Pero una lógica inductiva construida de acuerdo con la concepción de Carnap evitaría estas inconsistencias por otra razón también. Según esa concepción, y como ya señalamos en la sección 3, un razonamiento inductivo debe estar construido de tal modo que muestre que la información dada en los elementos de juicio, y que constituye las premisas, da un sustento inductivo más o menos fuerte a la conclusión; y concebida de esta manera, la inferencia inductiva no se presta al establecimiento categórico de la conclusión aunque se sepa o se admita que las premisas son enunciados verdaderos. Luego, la posibilidad de afirmar o aceptar enunciados incompatibles como resultado de inferencias inductivas no se plantea.

Pero en una lógica inductiva concebida de este modo, ¿no aparecerá en una forma diferente la dificultad planteada por las inconsistencias? Por ejemplo, si la información sobre la cantidad de esmeraldas observadas hasta ahora las presenta todas como verdes y verdules al mismo tiempo, ¿no llevará a razonar que esta información confiere una elevada probabilidad a dos predicciones incompatibles: "la primera esmeralda que se

[41] Véase C. G. Hempel, "Studies in the Logic of Confirmation", *Mind*, 54: 1-26 y 97-121 (1945), en particular sección 5 del capítulo I.

[42] Véase N. Goodman, "A Query on Confirmation", *The Journal of Philosophy*, 43: 383-385 (1946); Carnap, "On the Application of Inductive Logic" sección 3, y la respuesta de Goodman, "On Infirmities of Confirmation Theory", *Philosophy and Phenomenological Research*, 8: 149-151 (1947).

examine después de t será verde" y "la primera esmeralda que se examine después de t será verdul"? Nuevamente, la respuesta es negativa; pues como consecuencia de los postulados básicos de la probabilidad inductiva, la suma de las probabilidades que un conjunto lógicamente consistente de enunciados —por ejemplo, los elementos de juicio totales en t— confiere a dos hipótesis lógicamente incompatibles, es a lo sumo 1; luego, si una de las probabilidades se acerca a 1, la otra debe acercarse a 0.

6. *Sobre las reglas de decisión y creencia racionales*

Las reglas de inducción elementales consideradas en la sección anterior conciben el razonamiento inductivo como que conduce a la adquisición de nuevos enunciados sobre la base de los dados. A este respecto, concuerdan con la concepción familiar, según la cual los procedimientos inductivos, tanto en el nivel del sentido común como en el nivel científico, llevan a la *aceptación* de ciertas hipótesis empíricas sobre la base de elementos de juicio que les dan un apoyo mayor o menor, pero no, por lo general, un apoyo concluyente. El conjunto del conocimiento científico en un momento dado estaría representado, pues, por el conjunto de todos los enunciados aceptados por la ciencia en ese momento. La pertenencia a este conjunto sólo se le concedería a una hipótesis, por bien confirmada que esté, con carácter provisional, es decir, se le puede retirar el privilegio si aparecen en el futuro elementos de juicio desfavorables a ella.

El rechazo de nuestras reglas de inducción elementales, pues, plantea naturalmente la cuestión de si hay o no alguna manera alternativa consistente de construir este concepto del conocimiento científico, y más específicamente la noción de reglas que autoricen la adición de hipótesis empíricas "nuevas" con suficiente apoyo al conjunto de las aceptadas previamente. Esta cuestión pertenece, obviamente, a lo que Carnap llama la metodología de la inducción: concierne a la aplicación de la lógica inductiva a la formación de creencias racionales. Parece interesante, pues, indagar si no podría abordarse esa cuestión como un caso especial de otro problema muy general de aplicación que ha recibido mucha atención en años recientes, a saber, el problema de formular reglas para la elección o la decisión racional frente a diversas alternativas: la aceptación de una hipótesis sería concebible entonces como un caso de elección teórica entre hipótesis alternativas.

El problema de las reglas de decisión racional ha sido abordado recientemente en la teoría estadística de la adopción de decisiones y de la teoría de juegos, que no usan el concepto de probabilidad inductiva, y también ha sido investigado desde el punto de vista de la lógica inductiva. Aquí me limitaré a hacer un breve examen del enfoque de Carnap de esta cuestión. En la suposición de que se dispone de un sistema de lógica inductiva en el sentido de Carnap, el problema de la elección racional puede plantearse de la siguiente forma esquemática: Un agente X tiene que elegir uno de n cursos de acción, A_1, A_2, . . ., A_n, que, sobre la base de

los elementos de juicio totales e, de que dispone, se excluyen mutuamente y en conjunto agotan todas las posibilidades que se le ofrecen. El agente considera un conjunto O_1, O_2 ..., O_n de "resultados" diferentes posibles que, sobre la base de e, son mutuamente excluyentes y en conjunto exhaustivos (o sea, e implica lógicamente que se producirá uno solo de estos resultados). Luego, para cada una de esas acciones, digamos A_j, y uno cualquiera de esos resultados, digamos O_k, el sistema dado de lógica inductiva determina una probabilidad de que, dado e, A_j conducirá al resultado O_k. En verdad, si a_j y o_k son enunciados que describen a A_j y O_k, respectivamente, la probabilidad está dada por $c(o_k, e \cdot a_j)$.

Cuál sea el curso de acción racional para x en las circunstancias dadas dependerá, por supuesto, de cuáles sean sus objetivos, o, para decirlo en términos más amplios, de cuál sea el valor o disvalor que asigne a los diversos resultados que podrían producirse como efecto de su acción. En muchos estudios teóricos sobre la adopción racional de decisiones, y en particular en el tratamiento de Carnap del problema, se supone que los valores y disvalores en cuestión pueden representarse por un concepto cuantitativo de utilidad, es decir, por una función u que asigne a cada resultado posible O_k un número real, $u(O_k)$ o brevemente u_k, que indique la utilidad del resultado O_k para X. La tarea de especificar criterios operacionales para este concepto de utilidad —es decir, de especificar métodos para medir las utilidades de los resultados posibles (que pueden ser muy complejos) para una persona determinada— plantea difíciles problemas, que han sido el objeto de mucha labor teórica y experimental en años recientes.[43] Pero en este contexto no necesitamos entrar en estas cuestiones.

El problema que debemos resolver exige la formulación de una regla general de decisión tal que, dado cualquier e y cualquier conjunto de A_j y O_k que satisfagan las condiciones mencionadas, y dadas también las utilidades asignadas a los O_k, la regla determinará cuáles de los cursos de acción disponibles es racional adoptar en las circunstancias dadas. Carnap adopta una regla que lleva al agente a elegir una acción que le ofrece la mayor esperanza de utilidad. El valor de esperanza, o estimación de probabilidad, de la utilidad asociada con la acción A_j está dada por la fórmula:

(6.1) $u' (A_j, e) = c(o_1, e \cdot a_j) \cdot u_1 + \ldots + c(o_m, e \cdot a_j) \cdot u_m$,

y la regla de Carnap puede enunciarse del siguiente modo:

(6.2) *Regla de maximización de la utilidad estimada:*

En las circunstancias especificadas, elíjase un curso de acción pa-

[43] Para mayores detalles y referencias bibliográficas véase, por ejemplo, Carnap, *Logical Foundations of Probability*, sección 51; J. von Neumann y O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, 2ª ed., 1947); L. J. Savage, *The Foundations of Statistics* (Nueva York, Wiley, 1954), cap. 5; R. D. Luce y H. Raiffa, *Games and Decisions* (Nueva York, Wiley, 1957), cap. 2; R. B. Braithwaite, *Scientific Explanation* (Cambridge University Press, 1953), cap. VII.

ra el cual la estimación de la utilidad resultante sea un máximo, es decir, no sea superada por las estimaciones de utilidad asociadas con cualquiera de los cursos de acción alternativos.[44]

En un intento por aplicar este máximo al problema de la aceptación de reglas para hipótesis científicas, supongamos ahora que un científico dispone del conjunto de todos los enunciados aceptados por la ciencia en ese momento, que podemos suponer expresado en la forma de una complicada oración e; que ha inventado o se le ha presentado un conjunto de n hipótesis h_1, h_2, ..., h_n las cuales, sobre la base e, son incompatibles dos a dos a la par que en conjunto agotan todas las posibilidades (esto es, e implica lógicamente la negación de la conjunción de dos cualesquiera de las hipótesis, así como la disyunción de todas ellas); y que debe elegir uno de los siguientes $n + 1$ cursos de acción: Aceptar h_1 y sumarla a e; ...; aceptar h_n y sumarla a e; no aceptar ninguna de las n hipótesis y, por ende, no modificar e. El problema consiste en elaborar una regla que permita determinar cuál de las opciones es racional. Evidentemente este enfoque del problema de las reglas para la aceptación racional inductiva no supone el tipo de concepción estrechamente inductivista de la investigación científica que, aunque casi carece de adeptos en la actualidad, ha sido convertido en un caballito de batalla por algunos autores que escriben sobre el procedimiento de la ciencia. Más específicamente, no pensamos en una regla que, dados algunos elementos de juicio empíricos, permita *inferir* inductivamente "la" hipótesis o teoría, o simplemente una hipótesis o teoría, que explique los elementos de juicio dados. Más bien, suponemos aquí que se han presentado varias hipótesis rivales. La invención de tales hipótesis exige, en general, inventiva científica y, en los casos importantes, un gran genio; no se la realiza mediante el uso de reglas mecánicas de inducción. El problema inductivo aquí considerado es, más bien, el de determinar, sobre la base de los elementos de juicio, que pueden incluir los resultados de extensos ensayos, cuál, si es que hay alguna, de las hipótesis propuestas debe aceptarse y, de este modo, agregarse al *corpus* del conocimiento científico.

Ahora bien, el principio de decisión de Carnap (y análogamente, también, cursos de acción como el del principio del minimax desarrollado en la teoría de juegos y de decisiones estadísticas)[45] exige como base de una decisión racional una especificación tanto de los elementos de juicio totales como de las utilidades asignadas a los diversos resultados posibles de las acciones consideradas. En nuestro caso, los resultados posibles pueden describirse así: aumentar e con h_1, donde h_1 es verdadera; aumentar e con h_1, donde h_1 es falsa; ...; aumentar e con h_n, donde h_n es verdadera; aumentar e con h_n, donde h_n es falsa; no modificar e. ¿Qué utilidades debemos asignar a estos resultados? Esto por lo menos es claro: las utilidades deben reflejar el valor o disvalor que los diferentes resultados tienen desde el punto de vista de la investigación científica pura, más que

[44] Véase Carnap, *Logical Foundations of Probability*, pág. 269.
[45] Véase Carnap, observaciones en la sección 98 de *Logical Foundations of Probability*, así como la bibliografía indicada en la nota 43.

las ventajas o desventajas prácticas que podrían resultar de la aplicación de una hipótesis aceptada, según que ésta sea verdadera o falsa. A los tipos de utilidades caracterizadas vagamente de este modo llamémoslos *utilidades puramente científicas y utilidades epistémicas.*

Si pensamos que la proverbial "búsqueda de la verdad" en la ciencia tiende a establecer un sistema maximal de enunciados verdaderos, podemos tratar, como primer paso, de medir la utilidad de agregar una hipótesis h a e en términos de la fuerza de esa parte de la información contenida en h que no está contenida en e y que, por ende, va más allá de lo establecido anteriormente. La nueva información contenida en h está expresada por la oración $h \vee -e$; pues h es esquivalente a $(h \vee e) \cdot (h \vee -e)$. La primera de las dos oraciones que forman la conjunción se sigue de h y de e y representa la información dada por h y por e; la segunda, se sigue de h y expresa parte de la información dada por h, pero no tiene ningún contenido común con e, puesto que su disyunción con e es una verdad lógica.

Para representar la cantidad o la fuerza de la información dada por una oración usamos el concepto de *medida de contenido* para las oraciones de un lenguaje L (adecuadamente formalizado). Entendemos por tal medida de contenido cualquier función m que asigne a cada oración o de L un número $m(o)$ tal que: (I) $m(o)$ sea un número del intervalo comprendido entre 0 y 1, incluidos los extremos; (II) $m(o) = 1$ si o es lógicamente falsa (contradictoria); (III) si o_1 y o_2 no tiene ningún punto en común —es decir, si la oración $o_1 \vee o_2$, que expresa su contenido común es una verdad lógica— entonces $m(o_1 \cdot o_2) = m(o_1) + m(o_2)$; (IV) si o_1 y o_2 son lógicamente equivalentes, entonces $m(o_1) = m(o_2)$. Las medidas de contenido de este género pueden construirse fácilmente en ciertos tipos de lenguajes formalizados.[46]

Supongamos ahora que m sea una medida de contenido para un lenguaje formalizado de la ciencia empírica. Podemos hacer ahora, a título de ensayo, que la utilidad de agregar h a e sea igual a $m(h \vee -e)$ si h es la verdadera, e igual a $-m(\mathrm{h} \vee -e)$ si h es falsa. Con mayor generalidad, tomando en cuenta el principio de la utilidad marginal decreciente, podríamos hacer que la utilidad de agregar h a e sea directamente proporcional a la cantidad de nueva información suministrada por h, o al valor negativo de esa cantidad, según h sea verdadera o falsa, e inversamente proporiconal a la cantidad de información ya contenida en e. Esto conduciría a la siguiente definición:

(6.3) *Medida de contenido relativo de la utilidad puramente científica*: La utilidad puramente científica de agregar h a e es $k.m(h \vee -e)/ /m(e)$ cuando h es verdadera, y el negativo de este valor cuando h es falsa; k es una constante positiva.

[46] Véanse ejemplos específicos en C. G. Hempel y P. Oppenheim, "Studies in the Logic of Explanation", *Philosophy of Science,* 15: 135-175 (1948), en particular las secciones 8 y 9 del capítulo I; y R. Carnap e Y. Bar-Hillel, "An Outline of Theory of Semantic Information", Massachusetts Institute of Technology, Research Laboratory of Electronics. Technical Report Nº 247 (1952). Como base, véase también Carnap, *Logical Foundations of Probability,* sección 73.

Puede demostrarse[47] que si se adopta esta medida de la utilidad, independientemente de cuál de las muchas funciones de medida posibles sea m, entonces el principio de Carnap de maximizar la utilidad estimada conduce a la siguiente regla de decisión para el caso, caracterizado antes, de una elección entre las $n + 1$ alternativas de aceptar $h_1, \ldots,$ aceptar h_n o no aceptar ninguna de las hipótesis alternativas:

(6.4) *Regla de aceptación basada en la medida de contenido relativo de la utilidad.*
De las n hipótesis, a lo sumo una puede tener una probabilidad, sobre la base de e, que exceda de ½; si hay una, aceptarla. En caso contrario, puede haber a lo sumo dos hipótesis con una probabilidad de ½; en este caso, aceptar una de ellas o, alternativamente, no aceptar ninguna de las n hipótesis. Finalmente, si cada una de las n hipótesis tiene una probabilidad de menos de ½ sobre la base de e, no aceptar ninguna de ellas. (En el primer caso, la utilidad estimada será positiva; en todos los otros casos, será cero.)

De este modo, si se concibe la utilidad epistémica a la manera de (6.3), entonces el principio general de Carnap de maximizar la utilidad estimada brinda una regla que hace depender la aceptación de una o de ninguna de las n hipótesis rivales exclusivamente de las probabilidades que estas hipótesis poseen sobre la base de los elementos de juicio totales, e. Esta regla no puede conducir a inconsistencias inductivas, ya que la hipótesis aceptada debe tener una probabilidad de al menos ½, sobre la base de los elementos de juicio totales, y, por ende, no puede ser incompatible con éstos; y debe recordarse que los elementos de juicio totales representan, en nuestro caso, el conjunto de *todos* los enunciados aceptados en la ciencia en ese momento. Sin embargo, la regla (6.4) es insatisfactoria; en particular, es demasiado indulgente para ser apropiada como regla general del procedimiento científico. No debe interpretarse esto, sin embargo, en el sentido de que prueba que la regla de Carnap para la elección racional no puede ser una razonable regla de aceptación para las hipótesis científicas: muy probablemente sea defectuosa nuestra tosca definición de utilidad epistémica.

En verdad, aparte de suministrar nueva información verdadera o falsa, la adición de una hipótesis h a e tienen otros aspectos que son de importancia para la ciencia pura y que deben ser tomados en cuenta en el intento por definir un concepto de utilidad puramente científica. Por ejemplo, si h tiene el carácter de una ley general o de un principio teórico, su poder explicativo con respecto a los datos de importancia incluidos en e influirán vigorosamente sobre la utilidad potencial de aceptar h. Un factor estrechamente relacionado con éste sería, sin duda, la ganancia en simplicidad lógica que se agregaría al sistema total de los enunciados aceptados como resultado de incorporar h a él. Si es menester tomar en cuenta factores como éstos, se les debe dar definiciones claras y precisas. En años

[47] La demostración, que aquí omitiremos, es una generalización del razonamiento utilizado para fundamentar un resultado más limitado en el capítulo XII.

recientes se han dado algunos pasos iniciales en tal sentido,[48] pero aún se necesita mucho esfuerzo para alcanzar una concepto general razonablemente apropiado de utilidad epistémica.

El enfoque esbozado del problema de las reglas de aceptación inductivas concibe la formación de creencias empíricas racionales y el establecimiento del conocimiento científico como que implica el uso de ciertos principios inductivos que, en condiciones específicas, autorizan a aceptar provisionalmente una hipótesis sobre la base de un conjunto dado de elementos de juicio totales más que a determinar simplemente su grado de confirmación. Como alternativa, sería interesante investigar las maneras posibles de concebir la lógica de la creencia racional y del conocimiento científico sin suponer reglas de aceptación. Los únicos principios inductivos que se aducirían en tal concepción serían, hablando en términos generales, de carácter estadístico. Por ejemplo, podrían ser reglas inductivas como las consideradas por Carnap y podrían tomar formas como las de (3.4) y (3.5);[49] o podrían tener un carácter diferente, quizá de acuerdo con teorías estadísticas de la decisión.

En efecto, se ha sostenido recientemente, sobre todo por autores que tratan de los procedimientos estadísticos de decisión y de la teoría de juegos, que no tiene sentido claro hablar de la aceptación de una hipótesis científica *per se*, sin especificar un curso de acción basado en ella; y en particular, que en teoría de la decisión lo que se llama la aceptación de una hipótesis dada equivale siempre a la adopción de determinado curso de acción. Según este punto de vista, sería necesario concebir la noción de conocimiento científico sin usar la idea de aceptación para nada; o a lo sumo, sería menester concebir la aceptación como un concepto pragmático que no tiene equivalente en la lógica de la ciencia. Sin embargo, como he tratado de mostrar en otra parte,[50] este punto de vista, aunque apoyado por argumentos muy plausibles, también presenta dificultades.[51]

En la actualidad, creo que no está resulta la cuestión de si la idea de aceptación inductiva de una hipótesis en la ciencia pura puede ser objeto de una concepción clara y metodológicamente aclaratoria, y, de manera correlativa, si hay buenas razones para mantener la familiar noción de

[48] Para una definición del poder explicativo de hipótesis expresables en ciertos tipos simples de lenguajes formalizados, véase Hempel y Oppenheim, *loc. cit.*, secciones 8 y 9. Sobre el tema de la simplicidad, en el sentido a que aquí aludimos, véase K. Popper, *The Logic of Scientific Discovery* (Londres, Hutchinson, 1959), cap. VII y *passim*; y véase también el lúcido examen e intento de aclaración de S. Barker, *Induction and Hypothesis*, donde se encontrarán nuevas referencias bibliográficas, en particular a la obra de Kemeny.

[49] A este respecto, véanse las observaciones de Carnap en *Logical Foundations of Probability*, pág. 206.

[50] Véase Hempel, "Deductive-Nomological *vs.* Statistical Explanation", sección 12.

[51] Las consideraciones que hemos esbozado arrojan dudas, creo, sobre la opinión de que la cuestión "de si *aceptar* una hipótesis, si *creer* en ella... es más fácil de responder que la cuestión de si *actuar* sobre ella". Esta opinión ha sido expuesta por R. Chisholm en su libro *Perceiving A Philosophical Study* (Cornell University Press, 1957), págs. 10-11 (bastardillas del autor). La parte I de este libro, titulada "La ética de la creencia", contiene muchas observaciones aclaratorias sobre algunos de los problemas que hemos examinado en este ensayo.

reglas de inducción científica que autoricen a la aceptación de una hipótesis sobre la base de apropiados elementos de juicio. Para la ulterior aclaración de estos problemas será necesario elaborar de manera más completa y precisa las concepicones alternativas del conocimiento científico brevemente consideradas en esta sección. Esto exige nuevos análisis filosóficos, en el sentido de Carnap, que tiendan a dar una elucidación lógica[52] del concepto central del problema.

[52] Véase Carnap, *Logical Foundations of Probability*, cap. I.

Capítulo III

La ciencia y los valores humanos [1]

1. *El problema*

A menudo se llama a nuestra época la era de la ciencia y de la tecnología científica, y ello con razón: los avances realizados durante los últimos siglos por las ciencias naturales y, más recientemente, por las disciplinas psicológicas y sociológicas, han ampliado enormemente nuestro conocimiento y profundizando nuestra comprensión del mundo en que vivimos y de nuestros congéneres; y la aplicación práctica de las concepciones científicas nos está proporcionando un control creciente de las fuerzas de la naturaleza y las mentes de los hombres. Como consecuencia de esto, nos hemos acostumbrado, no sólo a la idea de una tecnología físico-química y biológica basada en los resultados de las ciencias naturales, sino también al concepto y aun a la práctica de una tecnología psicológica y sociológica que utiliza las teorías y los métodos elaborados por la investigación de la conducta.

Este aumento del conocimiento científico y de sus aplicaciones ha reducido mucho la amenaza de algunos de los más viejos y formidables flagelos del hombre, entre ellos el hambre y la peste; ha elevado el nivel material de vida del hombre y ha puesto a su alcance la realización de visiones que hasta hace unas pocas décadas habrían parecido fantásticas, como la exploración activa del espacio interplanetario.

Pero al alcanzar estos resultados, la tecnología científica ha dado origen a una multitud de problemas nuevos y profundamente inquietantes. El control de la fisión nuclear no nos ha suministrado solamente la confortable perspectiva de disponer de una nueva reserva de energía, sino también la constante amenaza de la bomba atómica y del grave daño, para las generaciones presentes y futuras, proveniente de los subproductos radiactivos del proceso de fisión, aun en sus usos pacíficos. Y el progreso mismo del conocimiento y la tecnología biológicos y médicos, que ha reducido de manera tan notable la mortalidad infantil y aumentado la esperanza de vida del hombre en grandes regiones del globo, también ha contribuido en

[1] Este capítulo apareció en R. E. Spiller (comp.), *Social Control in a Free Society*. Filadelfia, University of Pennsylvania Press, 1960, págs. 59-64. Se lo reimprime aquí, con algunas supresiones y adiciones, con la amable autorización de la University of Pennsylvania Press.

forma importante a la amenaza de la "explosión demográfica", al rápido crecimiento de la población terrestre que estamos contemplando en la actualidad y que, también, es motivo de grave preocupación para todos aquellos que piensan en el bienestar de las generaciones futuras.

Evidentemente, los avances de la tecnología científica de la cual nos enorgullecemos y que han impreso sus características en todos los aspectos de esta "edad de la ciencia" han planteado al mismo tiempo muchos nuevos y graves problemas que exigen una solución urgente. Es muy natural que, en su deseo de hacer frente a estos nuevos problemas, el hombre se vuelva nuevamente a la ciencia y la tecnología científica en busca de ayuda. Pero un momento de reflexión nos mostrará que los problemas que es menester abordar no son simplemente tecnológicos, sino un intrincado complejo de problemas tecnológicos y morales. Tomemos el caso de la explosión demográfica, por ejemplo. Sin duda, plantea problemas tecnológicos específicos. Uno de ellos es la tarea de satisfacer al menos las necesidades básicas materiales de una población en rápido crecimiento por medio de recursos limitados; otro, es la cuestión de los medios por los cuales es posible mantener bajo control el crecimiento mismo de población. Sin embargo, estas cuestiones técnicas no agotan el problema. Pues, a fin de cuentas, aun ahora disponemos de varias maneras de contrarrestar el crecimiento demográfico. Pero algunos de estos métodos anticonceptivos han sido y continúan siendo objeto de una intensa controversia inspirada en razones morales y religiosas, lo cual muestra que la adecuada solución del problema no sólo exige el conocimiento de medios técnicos de control, sino también patrones para evaluar los medios alternativos a nuestra disposición; y este segundo requisito plantea, evidentemente, problemas morales.

No es necesario ampliar la lista de ejemplos: todo medio de control técnico que la ciencia nos brinda puede ser empleado de muy diferentes maneras, y una decisión acerca del uso que le daremos nos arrastra a cuestiones de evaluación moral. Y aquí surge un problema fundamental sobre el cual quiero ahora concentrar la atención: ¿es posible responder a tales problemas de evaluación por medio de los métodos objetivos de la ciencia empírica, que han logrado tanto éxito al darnos un conocimiento confiable y a menudo aplicable prácticamente de nuestro mundo? ¿Pueden servir esos métodos para establecer criterios objetivos de lo correcto y lo incorrecto y, de este modo, suministrar normas morales válidas para determinar la conducta apropiada en nuestros asuntos individuales y sociales?

2. *El testeo científico*

Abordemos esta cuestión considerando primero, aunque sea de manera breve y esquemática, cómo se llega a un conocimiento científico objetivo. Podemos dejar de lado la cuestión de los *caminos del descubrimiento,* esto es, el problema de cómo surge una nueva idea científica, de cómo se concibe una nueva hipótesis o teoría científica. Para nuestro fin basta

considerar los *modos* científicos de *convalidación*, es decir, la manera como la ciencia empírica examina una hipótesis nueva propuesta y determina si se la debe aceptar o rechazar. Usaré aquí la palabra "hipótesis" para referirme muy ampliamente a cualquier enunciado o conjunto de enunciados de las ciencias empíricas, sea que trate de un suceso particular, sea que se proponga establecer una ley general, sea que formule una teoría más o menos compleja.

Como es bien sabido, las ciencias empíricas determinan la aceptabilidad de una hipótesis propuesta por medio de tests adecuados. A veces, un test semejante sólo supone lo que podría llamarse la observación directa de hechos atinentes a la cuestión. Este procedimiento puede usarse, por ejemplo, al testear enunciados tales como "afuera está lloviendo", "todas las bolitas de esta urna son azules", "la aguja de este amperímetro se detendrá en el punto 6 de la escala", etc. En estos casos, unas pocas observaciones directas bastarán, por lo común, para decidir si la hipótesis que se investiga debe aceptarse como verdadera o rechazarse como falsa.

Pero la mayoría de las hipótesis importantes de la ciencias empíricas no pueden ser testeadas de esta manera simple. La observación directa no basta para decidir, por ejemplo, si deben aceptarse o rechazarse las hipótesis de que la Tierra es una esfera, de que las características hereditarias se transmiten por los genes, de que todas las lenguas indoeuropeas provienen de una antigua lengua común, de que la luz es un proceso ondulatorio electromagnético, etc. Con hipótesis como éstas, la ciencia recurre a métodos indirectos de testeo y convalidación. Si bien estos métodos varían mucho en los procedimientos de detalle, todos ellos tienen la misma estructura y esquema básicos. Ante todo, de la hipótesis en estudio se infieren otros enunciados que describen ciertos fenómenos directamente observables que deben aparecer en circunstancias especificables, si la hipótesis es verdadera. Luego, los enunciados inferidos se testean directamente, es decir, se verifica si los fenómenos especificados se producen en realidad. Finalmente, se acepta o se rechaza la hipótesis propuesta a la luz del resultado de estos tests. Por ejemplo, la hipótesis de que la Tierra tiene forma esférica no es directamente testable mediante la observación, pero nos permite inferir que un barco que se aleje del observador debe parecer hundirse gradualmente bajo el horizonte, que es posible circunnavegar la Tierra siguiendo una trayectoria recta, que las fotografías a gran altura deben mostrar la curvatura de la superficie de la Tierra, que ciertas mediciones geodésicas y astronómicas deben dar tales y cuales resultados, etc. Tales enunciados inferidos pueden ser testeados más o menos directamente y, a medida que se verefica un número y una variedad cada vez mayor de ellos, la hipótesis se confirma en medida cada vez mayor. Eventualmente, una hipótesis puede llegar a estar tan bien confirmada por los elementos de juicio disponibles que se la acepta como establecida más allá de toda duda razonable. Sin embargo, no es posible probar de manera completa y definitiva ninguna hipótesis científica; existe siempre al menos la posibilidad teórica de que se descubran nuevos elementos de juicio que entren en conflicto con los enunciados observacionales inferidos de la hipótesis y que, por ende, conduzcan a su rechazo. La historia de la ciencia registra

muchos ejemplos en los que una hipótesis antaño aceptada luego fue abandonada a la luz de elementos de juicio adversos.

3. *Juicios instrumentales de valor*

Abordaremos ahora la cuestión de si este método de test y convalidación puede ser usado para establecer juicios morales de valor y, en particular, juicios según los cuales un curso específico de acción es bueno, correcto, apropiado o mejor que otros cursos alternativos de acción, o si debemos —o no debemos— actuar de determinadas maneras.

A título de ejemplo, consideremos la opinión de que es bueno criar a los niños en libertad, y malo tratarlos de manera restrictiva. Podría parecer que, al menos en principio, puede confirmarse científicamente esta opinión mediante apropiadas investigaciones empíricas. Supongamos, por ejemplo, que una cuidadosa investigación demuestra: (1) que la crianza restrictiva tiende a generar resentimiento y agresión contra los padres y otras personas que ejercen autoridad educacional, lo cual provoca sentimientos de culpa y ansiedad, y el eventual embotamiento de la iniciativa y las potencialidades creadoras del niño; en cambio, (2) que la crianza en libertad evita estas consecuencias, facilita las relaciones interpersonales, estimula la iniciativa y la autoconfianza, y permite al niño desarrollar y gozar de sus potencialidades. Estos enunciados, en particular cuando se los amplía adecuadamente, caen dentro de la esfera de la investigación científica; y aunque nuestro conocimiento de la cuestión sea de hecho muy limitada, supongamos, para los fines de nuestra argumentación, que han recibido una fuerte confirmación por tests cuidadosos. ¿No habría entonces demostrado objetivamente la investigación científica que es mejor criar a los niños en libertad que de una manera restrictiva?

Un poco de reflexión mostrará que no es así. Lo que se habría establecido es más bien un enunciado condicional; a saber, que *si* queremos que nuestros niños sean felices, emocionalmente estables y creadores, y no almas roídas por la culpa y perturbadas, *entonces* es mejor educarlos en libertad que de un modo restrictivo. Un enunciado como éste es un *juicio de valor relativo* o *instrumental.* Por lo general, un juicio de valor relativo dice que cierto tipo de acción, *M*, es bueno (o es mejor que otro tipo alternativo M_1) *si* se quiere lograr un objetivo *O* determinado. Pero decir esto equivale a afirmar, o bien que, en las circunstancias dadas, el curso de acción *M* conducirá definitivamente (o probablemente) al logro de *O*, o bien que no seguir el curso de acción *M* conducirá definidamente (o probablemente) al no logro de *O*. En otras palabras, el juicio de valor instrumental afirma, o que *M* es un medio suficiente (definida o probablemente) para alcanzar el fin u objetivo *O*, o que es (definida o probablemente) un medio necesario para alcanzarlo. Así, se puede reformular un juicio instrumental de valor en la forma de un enunciado que exprese un tipo universal o probabilístico de relación entre medios y fines, y que no contenga ningún término del discurso moral —tal como "bueno", "mejor",

"se debe"— en absoluto. Ahora bien, un enunciado de este tipo es, sin duda, una afirmación empírica que puede someterse a test empírico.

4. *Juicios categóricos de valor*

Desgraciadamente, esto no resuelve de manera completa nuestro problema. En efecto, después de someter a test un juicio relativo de valor referente a un cierto objetivo *O* y, supongamos, de haberlo confirmado, queda en pie la cuestión de si debe buscarse el objetivo *O* o si no es mejor tender a algún otro fin alternativo. La ciencia empírica puede establecer el enunciado condicional —por ejemplo— de que si deseamos evitar un sufrimiento intolerable a una persona que tiene una dolencia incurable, entonces una gran dosis de morfina brinda un medio para lograrlo; pero también puede sugerir maneras de prolongar la vida del paciente, aunque también prolongue su sufrimiento. Esto deja sin resolver la cuestión de si es correcto dar la primacía al objetivo de evitar sufrimientos sin esperanza sobre el de preservar la vida humana. Y esta cuestión requiere, no un juicio de valor relativo, sino *absoluto* o *categórico*, según el cual cierto estado de cosas (que puede haber sido propuesto como objetivo o fin) es bueno o es mejor que otra alternativa especificada. ¿Pueden someterse a test empírico y recibir confirmación semejantes juicios de valor categóricos?

Consideremos, por ejemplo, la oración "matar es malo". Ella expresa un juicio categórico de valor que, por implicación, calificaría categóricamente la eutanasia como un mal. Es evidente que esa oración no expresa una aserción que pueda ser testeada directamente por la observación, pues no se propone describir un hecho observable. Pero, ¿puede ser testeada indirectamente, infiriendo de ella enunciados que digan que en condiciones de prueba especificadas se producirán tales y cuales hechos observables? Otra vez, la respuesta es claramente negativa. En verdad, la oración "matar es malo" no tiene la función de expresar una aserción que pueda ser calificada de verdadera o falsa, sino que sirve para expresar un patrón para la evaluación moral o como norma de conducta. Un juicio categórico de valor también puede tener otras funciones; por ejemplo, puede servir para transmitir la aprobación o desaprobación de cierto tipo de acción por parte de quien lo enuncia, o su adhesión a las normas de conducta expresadas por el juicio de valor. Pero está ausente de él todo contenido empírico descriptivo. A este respecto, una oración tal como "matar es malo" difiere totalmente, por ejemplo, de la frase "matar está condenado como un mal por muchas religiones", la cual expresa una aserción fáctica que puede someterse a test empírico.

No es posible, pues, someter los juicios categóricos de valor a un test científico ni dar de ellos confirmación o desconfirmación, pues no expresan aserciones, sino patrones o normas de conducta. Fue Max Weber, creo, quien expresó la misma idea al observar que la ciencia es como un mapa: puede decirnos cómo llegar a un lugar determinado, pero no adónde ir. Gunnar Myrdal, en su libro *An American Dilemma* (pág. 1052), subraya,

en un estudio similar, que "los estudios teóricos o fácticos por sí solos no pueden conducir lógicamente a una recomendación práctica. Sólo puede deducirse una conclusión práctica o evaluativa cuando hay al menos una evaluación entre las premisas".

Sin embargo, se han realizado muchos intentos de basar sistemas de normas morales en los hallazgos de las ciencias empíricas, y sería interesante examinar en detalle el razonamiento que subyace en esos procedimientos. Pero en este contexto, sólo tenemos espacio para hacer unas pocas observaciones breves sobre este tema.

Podría parecer promisorio, por ejemplo, deducir juicios de valor de los resultados de un estudio objetivo de las necesidades humanas. Pero no es posible obtener ninguna deducción coherente de este tipo. En efecto, este procedimiento presupondría que es correcto o bueno satisfacer necesidades humanas, y esta presuposición es en sí misma un juicio categórico de valor: desempeñaría el papel de una premisa evaluativa, en el sentido del enunciado de Myrdal. Además, puesto que hay una gran cantidad de necesidades diferentes, y en parte conflictuales, de individuos y grupos, necesitaríamos, no el principio general de que se deben satisfacer las necesisdades humanas, sino un conjunto detallado de reglas acerca del orden y el grado preferenciales en los que deben satisfacerse las necesidades humanas, y de cómo resolver las aspiraciones conflictuales. Así, la premisa evaluativa que exige esta empresa tendría que ser, en realidad, un complejo sistema de normas; por consiguiente, está fuera de cuestión una deducción de normas evaluativas a partir de un estudio fáctico de las necesidades.

Varios sistemas de ética han considerado la teoría de la evolución como su fundamento, pero se encuentran en serio conflicto unas con otras en lo que respecta a sus afirmaciones fundamentales. Algunas de las variantes principales han recibido un aclarador examen de G. G. Simpson, en su libro *The Meaning of Evolution.* Uno de esos tipos de ética, al que Simpson llama una "ética de diente y garra", glorifica la lucha por la existencia que lleve a la supervivencia de los más aptos. Otro exige el ajuste armonioso de grupos o individuos para reforzar la probabilidad de su supervivencia, mientras que otros sistemas, aún, sostienen como norma suprema el creciente agregado de unidades orgánicas a niveles superiores de organización, a veces con la implicación de que el bienestar del Estado debe colocarse por encima del de los individuos que pertenecen a él. Es obvio que estos principios antagónicos no pueden haber sido inferidos válidamente de la teoría de la evolución, a menos que la teoría misma sea contradictoria, lo cual no parece muy probable.

Pero si la ciencia no puede brindarnos juicios catogóricos de valor, ¿qué es lo que puede servir como fuente de evaluaciones incondicionales? Puede entenderse esta pregunta en un sentido pragmático, como referido a las fuentes de las cuales los seres humanos obtienen de hecho sus valores básicos; o se la puede entender como referida a un aspecto sistemático de evaluación, a saber, a la cuestión de si puede hallarse un sistema apropiado de valores básicos sobre el cual fundar todas las otras evaluaciones.

La cuestión pragmática cae dentro de la esfera de la ciencia empírica. Sin entrar en detalles, podemos decir que una persona adquiere sus valores

—tanto aquellos a los que manifiesta adherir como aquellos a los que realmente se ajusta— principalmente de la sociedad en que vive, y en particular de ciertos grupos influyentes a los que pertenece, tales como su familia, sus compañeros de escuela, sus asociados en el trabajo, su iglesia, su club, su sindicato y otros grupos. En realidad, estos valores pueden variar de un caso a otro, según en cuál de estos grupos predomine la situación en la cual dicha persona se encuentra. En general, pues, las evaluaciones básicas de una persona no son el resultado de un examen cuidadoso y de una apreciación crítica de las alternativas posibles en mayor medida de lo que lo es su credo religioso. La conformidad con las normas de ciertos grupos desempeña aquí un papel muy importante, y sólo en raras ocasiones se ponen seriamente en tela de juicio los valores básicos. En verdad, en muchas situaciones tomamos decisiones y actuamos irreflexiblemente, en un sentido aún más fuerte, a saber, sin realizar ningún intento por basar nuestras decisiones en algún conjunto de normas morales explícitas conscientemente adoptadas.

Ahora bien, podría sostenerse que esta respuesta a la interpretación pragmática de nuestra pregunta refleja una lamentable inclinación humana a la inercia moral e intelectual, pero que el aspecto realmente importante de la pregunta es el sistemático: si queremos justificar nuestras decisiones, necesitamos normas morales de conducta del tipo incondicional. Pero ¿cómo es posible establecer esas normas? Si la ciencia no puede suministrar juicios catogóricos de valor, ¿hay otras fuentes en las que se los pueda obtener? ¿No podríamos, por ejemplo, convalidar un sistema de juicios categóricos de valor señalando que representa las normas morales de la Biblia, del Corán o de algún pensador o líder social inspirado? Es evidente que este procedimiento debe fracasar, pues la información fáctica aquí aducida podría servir para convalidar los juicios de valor en cuestión sólo si usáramos, además, una presuposición evaluativa según la cual *deben* cumplirse las directivas que provienen de la fuente invocada. De este modo, para completar el proceso de justificar una decisión o un juicio moral dados, deben aceptarse ciertos juicios de valor sin justificación ulterior, así como la demostración de un teorema geométrico exige la aceptación sin prueba de ciertas proposiciones como postulados. La búsqueda de una justificación de *todas* nuestras evaluaciones pasa por alto esta característica básica de la lógica de la convalidación y la justificación. Sin embargo, los juicios de valor aceptados sin ulterior justificación en un contexto dado no tienen por qué ser aceptados de una vez para siempre, con el compromiso de no ponerlos nunca en tela de juicio. Desarrollaremos este punto en la sección final de este ensayo.

Al concluir esta etapa de nuestro análisis, casi no es necesario destacar que las ideas expuestas en las páginas precedentes no implican ni propugnan la anarquía moral; en particular, no implican que cualquier sistema de valores sea tan bueno o tan válido como cualquier otro, ni que cada uno deba adoptar los principios morales que más le convengan. En efecto, estos principios tienen el carácter de juicios categóricos de valor y, por lo tanto, no pueden ser implicados por las consideraciones precedentes,

que son puramente descriptivas de ciertos aspectos lógicos, psicológicos y sociales de la evaluación moral.

5. *La elección racional: componentes empíricos y evaluativos*

Para comprender mejor la relación de la investigación con la evaluación categórica, preguntémonos qué ayuda podríamos recibir, al abordar un problema moral, de la ciencia en un estado ideal tal como el representado por la concepción de Laplace de una inteligencia científica superior, llamada a veces el demonio de Laplace. Esta ficción fue utilizada por Laplace a principios del siglo XIX para dar una caracterización vívida de la idea del determinismo causal universal. Se concibe este demonio como un observador perfecto, capaz de discernir con velocidad y exactitud infinitas todo lo que sucede en el universo en un momento dado. Es también un teórico ideal que conoce todas las leyes de la naturaleza y las ha combinado en una fórmula universal. Finalmente, es un matemático perfecto que, por medio de esta fórmula universal, puede inferir, a partir del estado observado del universo en un momento dado, el estado total del universo en cualquier otro momento. Así, el pasado y el futuro están ante sus ojos. Sin duda, es difícil imaginar que la ciencia pueda lograr alguna vez un grado superior de perfección.

Supongamos, pues, que, enfrentados con una decisión moral, podemos apelar al demonio de Laplace como a un consejero. ¿Qué ayuda podríamos obtener de él? Supongamos que tenemos que elegir uno de varios cursos de acción alternativos, y que queremos saber cuál de ellos *debemos* seguir. El demonio nos podrá decir, entonces, para cualquier opción, cuáles serían sus consecuencias para el curso futuro del universo, hasta el más ínfimo detalle y en el espacio y el tiempo más remotos. Pero, después de hacer esto para cada uno de los cursos alternativos de acción, el demonio habrá terminado su tarea: nos habrá dado toda la información que una ciencia ideal podría suministrar en las circunstancias dadas. Sin embargo, no habría resuelto nuestro problema moral, pues esto exige una decisión acerca de cuál de los diversos conjuntos alternativos de consecuencias delineadas por el demonio es el mejor, cuál de ellos debemos provocar. Y el peso de esta decisión aun caerá sobre nuestros hombros: somos nosotros quienes tendríamos que comprometernos con un juicio incondicional de valor indicando uno de los conjuntos de consecuencias como superior a sus alternativos. Ni siquiera el demonio de Laplace, o la ciencia ideal que él representa, puede aliviarnos de esta responsabilidad.

Al trazar este cuadro del demonio de Laplace como consejero en la adopción de decisiones, hemos trampeado un poco, pues si el mundo fuera tan estrictamente determinista como supone la ficción de Laplace, entonces el demonio sabría de antemano cuál elección haríamos y podría desengañarnos de la idea de que se nos abren varios cursos de acción ante nosotros. Sea como fuere, la teoría física contemporánea ha arrojado considerables dudas sobre la concepción clásica del universo como un sistema estrictamente determinista: se supone ahora que las leyes fundamen-

tales de la naturaleza tienen un carácter estadístico o probabilístico, no un carácter estrictamente universal y determinista.

Pero sean cuales fueren la forma y el alcance de las leyes que rigen en nuestro universo, obviamente nunca alcanzaremos un estado perfecto de conocimiento en lo concerniente a él; enfrentados a la necesidad de hacer una opción, nunca tenemos más que un conocimiento muy incompleto de las leyes de la naturaleza y del estado del mundo en el momento en que debemos actuar. Debemos, pues, tomar siempre nuestras decisiones sobre la base de una información incompleta, situación que nos permite a lo sumo prever las consecuencias de las opciones alternativas con cierta probabilidad. La ciencia puede rendirnos un servicio indispensable al brindarnos una información cada vez más amplia y confiable atinente a nuestro propósito; pero, nuevamente, nos corresponde a nosotros *evaluar* los diversos conjuntos probables de consecuencias de las opciones alternativas en consideración. Y esto exige la adopción de normas evaluativas atinentes al caso, normas que no están determinadas objetivamente por los hechos empíricos.

Este punto fundamental se refleja también en las teorías matemáticas modernas de la adopción de decisiones. Uno de los objetivos de estas teorías es la formulación de reglas de decisión que determinen una elección óptima en situaciones en las que son posibles varios cursos de acción. Para la formulación de reglas de decisión, estas teorías exigen que se cumplan al menos dos condiciones: (1) debe suministrarse una información fáctica que especifique los cursos de acción disponibles e indique, para cada uno de ellos, sus diferentes resultados posibles, más, si es posible, las probabilidades de ellos; (2) debe haber una especificación de los valores —a menudo prosaicamente llamados utilidades— que se asignan a los diferentes resultados posibles. Sólo cuando se satisfacen estas especificaciones evaluativas tiene sentido preguntar cuál de las opciones posibles es la mejor, considerando los valores asignados a sus resultados posibles.

En la teoría matemática de la decisión se han propuesto varios criterios de elección óptima. En el caso de que se disponga de las probabilidades de los diferentes resultados de cada acción, un criterio típico califica una elección de óptima si la utilidad probabilísticamente esperable de su resultado es al menos tan grande como la de cualquier elección alternativa. Otras reglas, tales como los principios del maximin y maximax, suministran criterios que son aplicables aun cuando no se disponga de las probabilidades de los resultados. Pero, hecho interesante, los diversos criterios entran en conflicto unos con otros, en el sentido de que, para la misma situación, a menudo eligen como óptimas opciones diferentes.

Puede considerarse que las políticas expresadas por los criterios antagónicos reflejan diferentes actitudes hacia el mundo, diferentes grados de optimismo o pesimismo, de espíritu de aventura o de cautela. Puede decirse, pues, que el análisis ofrecido por los modelos matemáticos actuales indica dos puntos en los cuales la adopción de decisiones no sólo exige información fáctica, sino también una evaluación categórica, vale decir, en la asignación de utilidades a los diferentes resultados posibles y en la adopción de una entre muchas reglas de decisión o criterios de elección

óptima rivales. Se desarrolla este tema con mayor detalle en la sección 10.2 de esta obra.

6. *"Presuposiciones" valorativas de la ciencia*

En las tres secciones anteriores nos hemos ocupado principalmente de la cuestión de si la evaluación y la decisión presuponen la investigación científica y el conocimiento científico, y si es así, en qué medida. Este problema tiene una contraparte que merece alguna atención en un examen de la ciencia y la evaluación, a saber, el de si el conocimiento y el método científicos presuponen la evaluación.

La palabra "presuponen" puede ser entendida en una serie de sentidos diferentes que exigen aquí una consideración separada. Ante todo, cuando una persona decide dedicarse a la labor científica y no a alguna otra actividad, y también cuando un científico elige un tema particular de investigación, estas opciones estarán presumiblemente determinadas en gran medida por sus preferencias, es decir, dependerán de cómo valore la investigación científica en comparación con las alternativas que se le abren, y por la importancia que asigne a los problemas que se proponga investigar. En este sentido explicativo y casi causal puede decirse ciertamente que las *actividades* científicas de los seres humanos presuponen evaluaciones.

Pero se plantean problemas mucho más desconcertantes cuando inquirimos si el *corpus* del conocimiento científico presupone juicios de valor; ese *corpus* podría estar representado por un sistema de enunciados aceptados de acuerdo con las reglas de la investigación científica. Aquí "presuponer" debe ser entendido en un sentido sistemático. Se utiliza este sentido cuando decimos, por ejemplo, que el enunciado "el cuñado de Enrique es ingeniero" presupone que Enrique tiene esposa o una hermana: en este sentido, un enunciado presupone todo lo que puede inferirse lógicamente de él. Pero, como observamos antes, ningún conjunto de enunciados científicos implica lógicamente un juicio incondicional de valor; luego, en este sentido, el conocimiento científico no presupone la evaluación.

Pero hay también otro sentido lógico de "presuponer". Podríamos decir, por ejemplo, que en la geometría euclidiana el teorema de la suma de los ángulos de los triángulos presupone el postulado de las paralelas, en el sentido de que este postulado es una parte esencial de las suposiciones básicas a partir de las cuales se deduce el teorema. Ahora bien, normalmente no se convalidan las hipótesis y teorías de la ciencia por su deducción a partir de elementos de juicio favorables a ellas (aunque puede suceder que se establezca un enunciado científico, tal como una predicción, por deducción a partir de un conjunto de enunciados previamente establecido y más amplio); más bien, como dijimos en la sección 2, habitualmente se las acepta sobre la base de elementos de juicio que sólo les prestan un apoyo parcial o "inductivo". Pero de todos modos, cabe preguntarse si los enunciados que representan conocimiento científico presuponen una evaluación en el sentido de que las razones por las cuales se los aceptan incluyen, a veces o siempre, ciertos juicios incondicionales de

valor. Nuevamente la respuesta es negativa. Las razones por las cuales se aceptan o se rechazan las hipótesis científicas las suministran los elementos de juicio empíricos, que pueden incluir hallazgos observacionales tanto como leyes y teorías previamente establecidas, pero que no incluyen, por cierto, ningún juicio de valor. Supongamos, por ejemplo, que en apoyo de la hipótesis de que un cinturón de radiación de un tipo especificado rodea la Tierra, un científico adujera, primero, ciertos datos observacionales, obtenidos quizá mediante instrumentos llevados por cohetes; segundo, ciertas teorías aceptadas aducidas en la interpretación de esos datos; y finalmente, ciertos juicios de valor tales como "es bueno descubrir la verdad". Obviamente, los juicios de valor serían descartados por carecer de toda atingencia lógica a la hipótesis propuesta, ya que no pueden contribuir a sustentarla ni a desconfirmarla.

Pero la cuestión de si la ciencia presupone evaluaciones en un sentido lógico puede plantearse, y recientemente se la ha planteado, aun de otra manera, que alude más específicamente a las presuposiciones valorativas del *método* científico. En las consideraciones precedentes representamos el conocimiento científico por un sistema de enunciados que tienen suficiente apoyo de los elementos de juicio disponibles como para ser aceptados de acuerdo con los principios del testeo y de la convalidación científicos. Observamos que, por lo común, los elementos de juicio observacionales sobre cuya base se acepta una hipótesis científica están lejos de bastar para establecer esta hipótesis de manera concluyente. Por ejemplo, la ley de Galileo no sólo se refiere a los casos pasados de caída libre cerca de la Tierra, sino también a todos los casos futuros, y éstos, evidentemente, no se encuentran entre nuestros elementos de juicio presentes. Luego, se acepta la ley de Galileo, y cualquier otra ley de las ciencias empíricas, sobre la base de elementos de juicio incompletos. Tal aceptación lleva consigo el "riesgo inductivo" de que la presunta ley pueda no ser válida con toda generalidad y de que los elementos de juicio futuros puedan llevar a los científicos a modificarla o abandonarla.

Una enunciación precisa de esta concepción del conocimiento científico exigiría, entre otras cosas, la formulación de reglas de dos tipos: Primero, *reglas de confirmación*, que especifiquen cuáles tipos de elementos de juicio son confirmatorios y cuáles desconfirmatorios para una hipótesis dada. Quizá tendrían también que determinar un *grado* numérico de apoyo de los elementos de juicio (o confirmación, o probabilidad inductiva) que un conjunto dado de éstos puede conferir a una hipótesis propuesta. En segundo término, tendría que haber *reglas de aceptación*: éstas especificarían cuán fuerte tiene que ser el apoyo en elementos de juicio que debe tener una hipótesis para que ésta pueda ser aceptada en el sistema de conocimiento científico; o, con mayor generalidad, en qué condiciones debe aceptar la ciencia una hipótesis propuesta y en qué condiciones la debe rechazar sobre la base de un conjunto dado de elementos de juicio.

Los estudios recientes sobre inferencia inductiva y testeo estadístico han dedicado mucho esfuerzo a la formulación de reglas apropiadas de ambos tipos. En particular, en muchas de estas investigaciones se han

considerado las reglas como casos especiales de reglas de decisión del tipo mencionado en la sección anterior. Aquí las decisiones son aceptar o rechazar una hipótesis propuesta sobre la base de elementos de juicio dados. Como observamos antes, la formulación de reglas de decisión "adecuadas" exige, en todo caso, la especificación previa de valoraciones que puedan servir como normas de adecuación. Las valoraciones requeridas, como se recordará, conciernen a los diferentes resultados posibles de las elecciones sometidas a las reglas de decisión. Ahora bien, cuando se aplica una regla científica de aceptación a una hipótesis específica sobre la base de un conjunto dado de elementos de juicio, los "resultados" posibles de la decisión pueden dividirse en cuatro tipos principales: (1) se acepta la hipótesis (como presumiblemente verdadera) de acuerdo con la regla y de hecho es verdadera; (2) se rechaza la hipótesis (como presumiblemente falsa) de acuerdo con la regla y de hecho es falsa; (3) se acepta la hipótesis de acuerdo con la regla, pero de hecho es falsa; (4) se rechaza la hipótesis de acuerdo con la regla, pero de hecho es verdadera. La ciencia aspira a realizar los dos primeros casos: la posibilidad de los otros dos representa el riesgo inductivo que debe correr toda regla de aceptación. Y el problema de formular reglas adecuadas de aceptación y rechazo no tiene ningún significado claro si no se han establecido normas de adecuación asignando valores o disvalores definidos a los diferentes "resultados" posibles de la aceptación o el rechazo. Es éste el sentido en el cual el método de establecer hipótesis científicas "presupone" valoraciones: la justificación de las reglas de aceptación y rechazo exige una referencia a juicios de valor.

En los casos en que la hipótesis testeada, si se la acepta, debe convertirse en la base de un curso específico de acción, los resultados posibles pueden conducir al éxito o al fracaso de la aplicación práctica considerada. En estos casos, los valores y disvalores en juego pueden ser expresables en términos de ganancias o pérdidas monetarias, y para las situaciones de este género la teoría de las funciones de decisión ha elaborado varias reglas de decisión utilizables en contextos prácticos tales como el control industrial de calidad. Pero cuando se trata de elaborar reglas de decisión para la aceptación de hipótesis en la investigación científica pura, donde no se contempla ninguna aplicación práctica, la cuestión de cómo asignar valores a los cuatro tipos de resultados mencionados antes se hace considerablemente más problemática. Pero en general, parece claro que las normas que gobiernan los procedimientos inductivos de la ciencia pura reflejan el propósito de obtener cierto fin, que podría describirse —de manera un tanto vaga— como el logro de un *corpus* de información confiable, amplio y teóricamente sistematizado acerca del mundo. Obsérvese que si quisiéramos, en cambio, elaborar un sistema de creencias o una concepción del mundo que fuera emocionalmente tranquilizante o estéticamente satisfactorio para nosotros, entonces no sería en modo alguno razonable insistir, como hace la ciencia, en que debe haber un estrecho acuerdo entre las creencias que aceptamos y los elementos de juicio empíricos; y los patrones de testabilidad y confirmación objetivas por elementos de juicio públicamente determinables tendrían que ser reemplazados por

patrones de aceptación de un tipo muy diferente. Las normas de procedimiento deben elaborarse en cada caso tomando en consideración los objetivos que se desea alcanzar; su justificación debe ser relativa a esos objetivos y, en este sentido, deben presuponerlos.

7. *Comparaciones finales*

Si, como hemos afirmado en la sección 4, la ciencia no puede suministrar una convalidación de juicios categóricos de valor, ¿puede el método científico desempeñar algún papel en la clarificación y solución de problemas de valoración y decisión morales? La respuesta es enfáticamente afirmativa. Trataré de demostrarlo con una breve exposición de las principales contribuciones que puede hacer la ciencia en este contexto.

Ante todo, la ciencia puede brindar la información fáctica necesaria para la resolución de problemas morales. Siempre se necesita de tal información, pues sea cual fuere el sistema de valores morales que adoptemos —sea egoísta o altruista, hedonístico o utilitario, o de cualquier otro tipo— el curso específico de acción que nos prescribe en una situación dada depende de los datos acerca de tal situación; y son el conocimiento y la investigación científicos los que nos brindan la información fáctica que necesitamos para la aplicación de nuestras normas morales.

Más específicamente, se necesita información fáctica, por ejemplo, para saber (a) si un objetivo en vista puede ser alcanzado en una situación determinada; (b) si es posible alcanzarlo, por cuáles medios alternativos y con cuáles probabilidades; (c) qué efectos colaterales y consecuencias ulteriores puede tener la elección de un medio determinado, aparte de la probabilidad de permitir el logro del fin deseado; (d) si varios objetivos deseados son realizables conjuntamente o si son incompatibles, en el sentido de que la realización de algunos de ellos excluya definida o probablemente la realización de otros.

Al darnos de este modo información que es indispensable como base de la decisión racional y responsable, la investigación científica puede hacernos cambiar algunas de nuestras valoraciones. Si descubriéramos, por ejemplo, que un determinado tipo de objetivo que hasta ahora hemos valorado muy alto sólo puede alcanzarse al precio de efectos colaterales y consecuencias ulteriores seriamente indeseables, podríamos otorgar un lugar menos elevado a dicho objetivo. Así, una información científica más amplia puede llevarnos a cambiar nuestras valoraciones básicas, no "desconfirmándolas", por supuesto, sino provocando un cambio en nuestra apreciación total de los problemas en cuestión.

En segundo término, y de una manera totalmente diferente, la ciencia puede aclarar ciertos problemas de valoración mediante un estudio psicológico y sociológico objetivo de los factores que influyen sobre los valores que admita un individuo o un grupo, de las formas en que cambian tales adhesiones valorativas y, quizá, de la manera como un sistema dado de valores puede contribuir a la seguridad emocional de un individuo o a la estabilidad funcional de un grupo.

Por supuesto que los estudios psicológicos, antropológicos y sociológicos de la conducta valorativa no pueden "convalidar" ningún sistema de normas morales. Pero sus resultados pueden provocar cambios morales, ampliando nuestro horizonte, señalándonos alternativas en las que no habíamos pensado o que nuestro grupo no sostiene y, de este modo, previniéndonos contra el dogmatismo o el parroquialismo moral.

Por último, la comparación con ciertos aspectos fundamentales del conocimiento científico puede ayudar a aclarar otras cuestiones concernientes a la valoración.

Si admitimos que las hipótesis y teorías científicas son siempre pasibles de revisión a la luz de nuevos elementos de juicio empíricos, ¿no estamos obligados a suponer que hay otra clase de enunciados científicos que no pueden estar sujetos a duda y reconsideración, a saber, los enunciados que describen hallazgos experimentales que sirven para testear las teorías científicas? ¿Esos informes simples y llanos de lo que se ha observado directamente en el laboratorio o en la labor científica de campo, por ejemplo, no deben considerarse inmunes a toda revisión concebible, irrevocables, una vez que se los ha establecido por observación directa? Los informes sobre fenómenos observados directamente a menudo han sido considerados, en efecto, como el cimiento inconmovible de todas las hipótesis y teorías científicas. Sin embargo, esta concepción es insostenible, pues ni siquiera en este ámbito encontramos una certeza definitiva e indiscutible.

Ello se debe, ante todo, a que los relatos de lo observado directamente están sujetos a errores que pueden brotar de diversas fuentes fisiológicas y psicológicas. En verdad, a menudo es posible verificar la exactitud de una observación comparándola con los informes de otros observadores o con datos relacionados con ella obtenidos mediante algún procedimiento indirecto, como la película filmada del final de una carrera de caballos; y esa comparación puede llevar al rechazo de lo que se consideraba previamente como una descripción correcta de un fenómeno observado directamente. Hasta hay teorías que nos permiten explicar y prever algunos tipos de errores observacionales, y en tales casos no se vacila en poner en duda y rechazar ciertos enunciados que se proponen simplemente registrar lo que se ha observado directamente.

A veces, hallazgos experimentales relativamente aislados pueden entrar en conflicto con una teoría que tiene un fuerte apoyo de un gran número y una gran variedad de otros datos. En este caso, bien puede suceder que se rechace la admisión al sistema de enunciados científicos aceptados a esa parte de los datos antagónicos, no a la teoría, aunque no se tenga ninguna explicación del presunto error de observación. En casos como éstos, no es el hallazgo observacional aislado el que decide si la teoría debe mantenerse, sino que es la teoría previamente bien fundamentada la que determina si un informe observacional debe considerarse como descripción de un hecho empírico real. Por ejemplo, el informe de que durante una sesión espiritista un mueble flotó libremente separado del suelo, normalmente será rechazado porque entra en conflicto con principios físicos muy bien confirmados, aun en ausencia de una explicación

específica del informe, por ejemplo, en términos de un engaño deliberado por parte de la médium o de una elevada sugestionabilidad por parte del observador. Análogamente, los hallazgos experimentales comunicados por el físico Ehrenhaft, que pretendía refutar el principio de que todas las cargas eléctricas son múltiplos enteros de la carga del electrón, no llevaron al abandono de este principio —ni siquiera a introducir en él una ligera modificación— el cual forma parte de una teoría que tiene un sustento experimental sumamente fuerte y variado. Es innecesario decir que ese rechazo de informes observacionales en razón de su antagonismo con teorías bien establecidas exige considerable cautela; de lo contrario, podría usarse una teoría, una vez aceptada, para rechazar todo elemento de juicio adverso que pudiera hallarse, procedimiento dogmático totalmente irreconciliable con los objetivos y el espíritu de la indagación científica.

Entonces, ni siquiera son irrevocables los informes sobre fenómenos observados directamente; éstos no ofrecen un fundamento inconmovible para todo el sistema del conocimiento científico. Pero esto no excluye en modo alguno la posibilidad de testear las teorías científicas con referencia a datos obtenidos por observación directa. Como hemos señalado, los resultados obtenidos mediante tal verificación directa no pueden considerarse como absolutamente indiscutibles e irrevocables; ellos mismos son pasibles de otros tests, que pueden realizarse si hay motivo de duda. Pero obviamente, para elaborar cualquier creencia acerca del mundo, para aceptar o rechazar —aunque sea provisoriamente— una hipótesis o una teoría, debemos detener el proceso de testeo en alguna parte; debemos aceptar algunos enunciados que expresen elementos de juicio y considerarlos suficientemente dignos de crédito como para no requerir más investigación por el momento. Y sobre la base de tales elementos de juicio podemos decidir qué crédito dar a la hipótesis testeada y si aceptarla o rechazarla.

Me parece que este aspecto de la investigación científica guarda cierto paralelismo con el caso de la valoración juiciosa y la decisión racional. Para hacer una elección racional entre diversos cursos de acción, debemos considerar, ante todo, qué consecuencias probablemente tenga cada una de las elecciones alternativas. Esto suministra una base para hacer ciertos juicios relativos de valor que son atinentes a nuestro problema. Para lograr *este* conjunto de resultados, debe seguirse este curso de acción; para alcanzar *ese otro* conjunto de resultados, debe elegirse tal o cual curso de acción, etcétera. Pero para llegar a una decisión, aún tenemos que establecer los valores relativos de los conjuntos alternativos de consecuencia que podemos obtener; y esto, como ya señalamos, exige la aceptación de un juicio incondicional de valor, que determinará entonces nuestra elección. Pero no es necesario considerar esta aceptación como definitiva e irrevocable, como obligatoria siempre para todas nuestras decisiones futuras: un juicio incondicional de valor, aunque aceptado una vez, queda sujeto a reconsideración y cambio. Supongamos, por ejemplo, que debemos elegir, como votantes o como miembros de la administración municipal, entre varias políticas sociales alternativas, algu-

nas de las cuales se proponen mejorar ciertas condiciones materiales de vida, mientras que otras aspiran a satisfacer necesidades culturales de diversos tipos. Para llegar a una decisión, debemos comprometernos con la asignación de un valor superior a uno u otro de esos objetivos. Pero aunque el juicio aceptado sirva como juicio incondicional y básico de valor para la decisión en vista, no por ello nos comprometemos con ella para siempre; bien podemos reconsiderar nuestras normas e invertir nuestro juicio más tarde; y aunque esto no puede anular la decisión anterior, conducirá a decisiones diferentes en el futuro. Así, para llegar a una decisión concerniente a un problema moral, debemos aceptar algunos juicios incondicionales de valor; pero no es necesario considerar éstos como definitivos, en el sentido absoluto de ser para siempre la norma para todas nuestras decisiones, de igual modo que no es necesario considerar los enunciados que expresan elementos de juicio y que son la base del testeo de hipótesis científicas como irrevocables para siempre. Todo lo que se necesita en ambos contextos son definitivos *relativos*, por así decir: un conjunto de juicios —morales o descriptivos— que sean aceptados por el momento sin necesidad de un examen ulterior. Estos juicios definitivos relativos nos permiten mantener la mente abierta a la posibilidad de efectuar cambios en nuestros compromisos y creencias hasta ahora indiscutidos; y la experiencia sugiere, sin duda, que para responder a las exigencias del presente y del futuro, deberemos más que nunca alejarnos del dogmatismo y mantener una mente crítica y abierta.

Segunda Parte

CONCEPCIONES DE SIGNIFICACION COGNOSCITIVA

Capítulo IV

Criterios empiristas de significación cognoscitiva: problema y cambios[1]

1. *La concepción empirista general de la significación cognoscitiva y empírica*

Un principio básico del empirismo contemporáneo es que una oración hace una afirmación cognoscitivamente significante, y —por ende— puede decirse de ella que es verdadera o falsa, si y sólo si: (1) es analítica o contradictoria, en cuyo caso se dice que tiene un sentido o significado puramente lógico; o (2) si es pasible, al menos potencialmente, de testeo por elementos de juicio experienciales, en cuyo caso se dice que tiene sentido o significado empírico. El supuesto básico de este principio, y en particular de su segunda parte, el llamado criterio testabilista del sentido (o mejor, de la posesión de sentido), no es exclusivo del empirismo: también es característico del operacionalismo contemporáneo y, en cierto sentido, del pragmatismo; en efecto, el principio pragmatista de que una diferencia debe dar origen a una diferencia para ser una diferencia, puede interpretarse en el sentido de que una diferencia verbal entre dos oraciones debe dar origen a una diferencia en las implicaciones experienciales para reflejar una diferencia de significado.

Es demasiado conocido para que sea menester recordarlo aquí cómo esta concepción general del discurso cognoscitivamente significante condujo al rechazo, por considerarlas carentes de significado lógico y empírico, de varias formulaciones de la metafísica especulativa y hasta de ciertas hipótesis de las ciencias empíricas. Creo que el propósito general del criterio empirista del significado es básicamente correcto y que, a pesar de las excesivas simplificaciones en su uso, su aplicación crítica ha sido, en conjunto, aclaradora y saludable. Pero me siento menos confiado en la posibilidad de reformular la idea general en la forma de criterios precisos y generales que establezcan una clara línea demarcatoria

[1] Con algunas omisiones y otros cambios, este capítulo combina los contenidos de dos artículos: "Problems and Changes in the Empiricist Criterion of Meaning", *Revue Internationale de Philosophie* Nº 11, págs. 41-63 (enero de 1950) y "The Concepts of Cognitive Significance: A Reconsideration", *Proceedings of the American Academy of Arts and Science* 80, Nº 1, págs. 61-77 (1951). Se reimprime este material con la amable autorización del director de la *Revue Internationale de Philosophie* y de la American Academy of Arts and Sciences.

(a) entre los enunciados de significación puramente lógica y los enunciados de significación puramente empírica, y (b) entre las oraciones que tienen significación cognoscitiva y las que no la tienen.

En este capítulo me propongo reexaminar estas distinciones, tal como las concibe el empirismo actual, y señalar algunas de las dificultades que presentan. El análisis se ocupará principalmente de la segunda de las dos distinciones. Con respecto a la primera, me limitaré a algunas observaciones breves.

2. *Los primeros criterios testabilistas del significado y sus inconvenientes*

Observemos primero que todo criterio general de significación cognoscitiva deberá satisfacer ciertos requisitos para ser aceptable. De ellos, destacaremos uno que consideraremos aquí como expresión de una *condición de adecuación* necesaria, aunque en modo alguno suficiente, de los criterios de significación cognoscitiva.

(A) Si de acuerdo con un criterio dado de significación cognoscitiva una oración N carece de significado, entonces también deben carecer de él todas las oraciones compuestas en las que N aparezca de manera no vacía como componente. En efecto, si no puede asignarse a N significativamente un valor de verdad, entonces es imposible asignar valores de verdad a las oraciones compuestas que contienen a N; por consiguiente, también se las debe considerar como carentes de significado.

Del requisito (A) se desprenden dos corolarios:

(A1) Si de acuerdo con un criterio dado de significación cognoscitiva una oración S carece de significado, entonces también carece de él su negación, $\sim S$.

(A2) Si de acuerdo con un criterio dado de significación cognoscitiva una oración N carece de significado, entonces también carece de él toda conjunción $N \cdot S$ y toda disyunción $N \vee S$, tenga o no S significado de acuerdo con el criterio dado.

Pasemos ahora a los intentos iniciales realizados por el empirismo actual para establecer criterios generales de significación cognoscitiva. Esos intentos se rigieron por la idea de que una oración, para hacer una afirmación empírica, debe apoyarse o entrar en conflicto con fenómenos que potencialmente son pasibles de observación directa. A las oraciones que describen tales fenómenos potencialmente observables, se produzcan o no, se las puede llamar oraciones observacionales. Más específicamente, una *oración observacional* puede ser concebida como una oración, sea verdadera o falsa, que se afirma o niega que un objeto, o un grupo de objetos, específico y de tamaño macroscópico tiene una *característica observable* particular, es decir, una característica cuya presencia o ausencia puede discernirse, en circunstancias favorables, por observación directa.[2]

[2] Las oraciones observacionales de este tipo pertenecen a lo que Carnap ha llamado el lenguaje de objeto; véase, por ejemplo (1938), págs. 52-53. El hecho de que sean adecuadas para formular los datos que sirven de base a los tests empíricos se hace evidente, en particular, por los procedimientos de testeo intersubjetivo utilizado en la ciencia y en grandes regiones de la investigación empírica en el nivel

La tarea de establecer criterios de significación empírica se transforma, así, en el problema de caracterizar de una manera precisa la relación que existe entre una hipótesis y una o más oraciones observacionales cuando los fenómenos que estas últimas describen confirman o desconfirman la hipótesis. La capacidad de una oración dada de entrar en esa relación con algún conjunto de oraciones observacionales caracterizaría, pues, su testabilidad-en-principio y por ende su significación empírica. Examinemos ahora brevemente los intentos principales que se han hecho para obtener criterios de significación de este modo.

Uno de los primeros criterios se expresaba en el llamado *requisito de verificabilidad*. Según él, una oración es empíricamente significativa si y sólo si no es analítica y es susceptible, al menos en principio, de una verificación completa mediante elementos de juicio observacionales; es decir, si es posible describir elementos de juicio observacionales que, si se los obtuviera realmente, establecerían de manera concluyente la verdad de la oración.[3] Con ayuda del concepto de oración observacional podemos reformular este requisito del siguiente modo: una oración S tiene signi-

del sentido común. En las discusiones epistemológicas, se supone con frecuencia que los elementos de juicio finales para las creencias acerca de asuntos empíricos consisten en percepciones y sensaciones cuya descripción exige un tipo de lenguaje fenomenalista. No podemos discutir aquí los problemas específicos vinculados con el enfoque fenomenalista, pero debemos mencionar que, sea como fuere, todas las consideraciones críticas presentadas en este capítulo con respecto al criterio de testabilidad son aplicables, *mutatis mutandis*, al caso de una base fenomenalista.

[3] Originalmente, se entendía que los elementos de juicio admisibles se limitaban a lo que es observable por el parlante y, quizá, por sus colegas durante sus vidas. Concebido de este modo, el criterio excluye, por considerarlos cognoscitivamente carentes de significado, todos los enunciados acerca del futuro distante y el pasado remoto, como han señalado, entre otros, Ayer (1946), capítulo I, Pap (1949), capítulo 13, en particular págs. 333 y sigs., y Russell (1948), págs. 445-447. Se evita esta dificultad, sin embargo, si admitimos que los elementos de juicio pueden consistir en un conjunto finito de "datos observacionales lógicamente posibles", cada uno de ellos formulado en una oración observacional. Así, por ejemplo, la oración S_1, "la lengua del mayor dinosaurio del Museo de Historia Natural de Nueva York era azul o negra" es completamente verificable, en nuestro sentido, pues es una consecuencia lógica de la oración S_2, "la lengua del mayor dinosaurio del Museo de Historia Natural de Nueva York era azul", y ésta es una oración observacional en el sentido indicado.

Y si se conciben el concepto de *verificabilidad en principio* y el concepto más general de *confirmabilidad en principio*, que serán considerados más adelante, como referentes a *elementos de juicio lógicamente posibles* expresados por oraciones observacionales, entonces, de modo análogo, se desprende que la clase de los enunciados verificables, o al menos confirmables, incluye en principio aserciones tales como que el planeta Neptuno y el continente antártico existieron antes de ser descubiertos, y que la guerra atómica, si no se la evita, conducirá al exterminio del planeta. Las objeciones que plantea Russell (1948), págs. 445 y 447, contra el criterio de verificabilidad con referencia a estos ejemplos, no se aplican, pues, a ellos si se entiende el criterio del modo sugerido. Digamos de paso que los enunciados del tipo mencionado por Russell, que no son en realidad verificables por ningún ser humano, ya fueron reconocidos en forma explícita como cognoscitivamente significantes por Schlick (1936), parte V, quien argüía que la imposibilidad de verificarlos es "meramente empírica". La caracterización de la verificabilidad con ayuda del concepto de oración observacional, tal como aquí se sugiere, podría servir para enunciar de modo más explícito y riguroso esa idea.

ficado empírico si y sólo si es posible indicar un conjunto finito de oraciones observacionales, $O_1, O_2, \ldots, O_n$, tales que, si fuesen verdaderas, también S lo sería necesariamente. Formulada de esta manera, sin embargo. esta condición también se satisface si S es una oración analítica o si las oraciones observacionales dadas son lógicamente incompatibles entre sí. Mediante la siguiente formulación excluimos estos casos y, al mismo tiempo, expresamos dicho criterio de modo más preciso:

(2.1) *Requisito de la completa verificabilidad en principio.* Una oración tiene significado empírico si y sólo si no es analítica y se deduce lógicamente de alguna clase finita y lógicamente consistente de oraciones observacionales.[4] No es necesario que estas oraciones observacionales sean verdaderas, pues lo que el criterio debe dilucidar es la testabilidad por "fenómenos potencialmente observables" o la testabilidad "en principio".

De acuerdo con el concepto general de significación cognoscitiva esbozado antes, una oración será clasificada ahora como cognoscitivamente significante si es analítica o contradictoria, o si satisface el requisito de verificabilidad.

Este criterio, sin embargo, tiene algunos defectos serios. Uno de ellos ha sido señalado por varios autores:

a) Supongamos que las propiedades de ser una cigüeña y de tener patas rojas son ambas propiedades observables, y que la primera no implica lógicamente la segunda. Entonces, la oración:

($S1$) Todas las cigüeñas tienen patas rojas

no es analítica ni contradictoria; y evidentemente no es deducible de un

[4] Como se ha subrayado con frecuencia en la bibliografía empirista, el término "verificabilidad" debe indicar, por supuesto, la posibilidad de concebir o, mejor dicho, la posibilidad lógica de elementos de juicio de tipo observacional que, si se hallara realmente, constituirían elementos de juicio concluyentes con respecto a la oración dada. Pero no pretende establecer la posibilidad técnica de realizar los tests necesarios para obtener tales elementos de juicio, y menos aún la posibilidad de hallar fenómenos directamente observables que constituyan elementos de juicio concluyentes acerca de esa oración, que equivaldría a la existencia real de tales elementos de juicio y, por ende, implicaría la verdad de la oración dada. Observaciones análogas se aplican a los términos "refutabilidad" y "confirmabilidad". Es evidente que este punto ha sido pasado por alto en algunos exámenes críticos del criterio de verificabilidad. Así, por ejemplo, Russell (1948), pág. 448, concibe la verificabilidad como la existencia real de un conjunto de sucesos que constituyen una verificación concluyente. Esta concepción, que nunca ha sido propugnada por ningún empirista lógico, debe ser —naturalmente— inadecuada, puesto que de acuerdo con ella no podría establecerse la significación empírica de una oración sin reunir elementos de juicio empíricos, en cantidad suficiente —además— como para permitir elaborar una prueba concluyente de la oración en cuestión. No cabe sorprenderse, pues, de que esta inusitada interpretación de la verificación lleve a Russell a la siguiente conclusión: "De hecho, que una proposición sea verificable no es en sí mismo verificable" (*loc. cit.*). En realidad, de acuerdo con la interpretación empirista de la verificabilidad completa, todo enunciado que afirme la verificabilidad de una oración S cuyo texto se cita, es analítica o contradictoria, pues la decisión de si existe una clase de oraciones observacionales que impliquen S, es decir, de si es posible formular tales oraciones observacionales, sean verdaderas o falsas, esa decisión, pues, es una cuestión puramente lógica.

conjunto finito de oraciones observacionales. Luego, de acuerdo con el criterio considerado, *S*1 está desprovista de significación empírica; y por lo tanto, también carecen de ella todas las otras oraciones que expresan regularidades universales o leyes generales. Y puesto que las oraciones de este tipo constituyen parte de teorías científicas, debe considerarse que el requisito de verificabilidad es demasiado restrictivo a este respecto.

De igual modo, el criterio descalifica a todas las oraciones tales como "para toda sustancia existe algún solvente", que contiene tanto cuantificadores universales como existenciales (esto es, los términos "todo" y "algunos" o sus equivalentes), pues no es posible deducir lógicamente ninguna oración de este tipo de un conjunto finito de oraciones observacionales.

Hay otros dos defectos del requisito de verificabilidad que no parecen haber recibido mucha atención:

b) Como se ve fácilmente, la negación de *S*1

(~ *S*1) Existe al menos una cigüeña que no tiene patas rojas,

es deducible de dos oraciones observacionales cualesquiera del tipo "a es una cigüeña" y "a no tiene patas rojas". Luego ~ *S*1 es cognoscitivamente significante según nuestro criterio, pero *S*1 no lo es, y esto constituye una violación de la condición (*A*1).

c) Sea *S* una oración que satisface el requisito de verificabilidad y *N* una oración que no lo satisface. Entonces, *S* es deducible de algún conjunto de oraciones observacionales; por consiguiente, por una conocida regla de la lógica, *S* v *N* es deducible del mismo conjunto y, por ende, es cognoscitivamente significante de acuerdo con nuestro criterio. Esto viola la condición (*A*2)[5]

Consideraciones estrictamente similares se aplican a un criterio alternativo que hace de la completa refutabilidad en principio la característica definitoria de la significación empírica. Formulamos este criterio del siguiente modo:

(2.2) *Requisito de la refutabilidad completa en principio.* Una oración tiene significado empírico si y sólo si su negación no es analítica y se sigue lógicamente de alguna clase finita y lógicamente consistente de oraciones observacionales.

[5] Los argumentos aquí aducidos contra el criterio de verificabilidad también prueban lo inadecuado de una concepción estrechamente relacionada con ella, a saber, que dos oraciones tienen la misma significación cognoscitiva si cualquier conjunto de oraciones observacionales que verifique a una de ellas verifica también a la otra, y recíprocamente. Así, por ejemplo, según este criterio, a dos leyes generales cualesquiera habría que asignarles la misma significación cognoscitiva, pues ningún conjunto de oraciones observacionales confirma ninguna ley general. Debe distinguirse claramente, la concepción a la que acabamos de aludir, de una posición que Russell examina en su análisis crítico del criterio positivista del significado. Es "la teoría de que dos proposiciones cuyas consecuencias verificadas son idénticas tienen la misma significación" (1948), pág. 448. Esta tesis es insostenible, en verdad, pues las consecuencias de un enunciado que hayan sido verificadas en un momento dado son, como es obvio, una cuestión de accidente histórico que no puede servir para establecer la identidad de significación cognoscitiva. Pero no sé de ningún empirista lógico que haya sostenido alguna vez tal "teoría".

Este criterio califica a una oración de empíricamente significativa si su negación satisface el requisito de la verificabilidad completa; como era de esperar, por consiguiente, es inadecuado por las mismas razones que el anterior:

a) Niega significación cognoscitiva a hipótesis puramente existenciales, tales como "existe al menos un unicornio", y a todas las oraciones cuya formulación exige una cuantificación mixta —es decir, universal y existencial—, por ejemplo, "para todo compuesto existe algún solvente", pues ni unas ni otras pueden ser refutadas de modo concluyente por un número finito de oraciones observacionales.

b) Si "*P*" es un predicado de observación, entonces la aserción de que todas las cosas tienen la propiedad *P* es calificada de significante; pero su negación, por ser equivalente a una hipótesis puramente existencial queda descalificada [compárese con (a)]. Luego, el criterio (2.2) da origen al mismo dilema que (2.1).

c) Si una oración *S* es completamente refutable mientras que *N* es una oración que lo no es, entonces su conjunción *S* · *N* (esto es, la expresión obtenida vinculando las dos oraciones mediante la palabra "y") es completamente refutable, pues si la negación de *S* es implicada por una clase de oraciones observacionales, entonces la negación de *S* · *N* es, *a fortiori*, implicada por la misma clase. Así, el criterio asigna significación empírica a muchas oraciones que un criterio empirista adecuado excluiría, por ejemplo, a "todos los cisnes son blancos y lo absoluto es perfecto".

En suma, pues, las interpretaciones del criterio de testabilidad en términos de la verificabilidad completa o de la refutabilidad completa son inadecuadas porque son demasiado restrictivas en un sentido y demasiado inclusivas en otro, y porque ambas violan el requisito fundamental A.

Se han hecho varias tentativas para evitar estas dificultades construyendo el criterio testabilidad de modo que se exija solamente una confirmabilidad parcial y quizás indirecta de las hipótesis empíricas por elementos de juicio observacionales.

Una formulación sugerida por Ayer [6] es característica de esas tentativas por establecer un criterio claro y suficientemente amplio de confirmabilidad. Dice, en efecto, que una oración *S* tiene contenido empírico si es posible deducir de *S*, junto con apropiadas hipótesis subsidiarias, oraciones observacionales que no son deducibles de las hipótesis subsidiarias solamente.

Esta condición se inspira en una consideración más minuciosa de la estructura lógica del testeo científico; pero es demasiado liberal, tal como está. En realidad, como el mismo Ayer ha señalado en la segunda edición de su libro *Language, Truth and Logic*,[7] su criterio atribuye contenido empírico a cualquier oración. Así por ejemplo, si *S* es la oración "lo absoluto es perfecto", basta elegir como hipótesis subsidiaria la ora-

[6] (1936, 1946), cap. I. La argumentación contra los requisitos de verificabilidad y refutabilidad, y a favor de un requisito de la confirmabilidad y desconfirmabilidad parcial, se halla muy claramente expuesta por Pap (1949), capítulo 13.

[7] (1946), 2ª edición, págs. 11-12.

ción "si lo absoluto es perfecto, entonces esta manzana es roja" para permitir la deducción de la oración observacional "esta manzana es roja", que evidentemente no se desprende de la hipótesis subsidiaria solamente.

Para hacer frente a esta objeción, Ayer propuso una versión modificada de su criterio de testabilidad. En efecto, la modificación restringe las hipótesis subsidiarias mencionadas en la versión anterior a oraciones que son analíticas o de las que puede mostrarse independientemente que son testeables en el sentido del criterio modificado.[8]

Pero puede mostrarse fácilmente que este nuevo criterio, como el requisito de la refutabilidad completa, asigna significación empírica a toda conjunción $S \cdot N$, donde S satisface el criterio de Ayer mientras que N es una oración tal como "el absoluto es perfecto", que debería ser descalificada por dicho criterio. En verdad, cualquier consecuencia que pueda ser deducida de S con ayuda de hipótesis subsidiarias permisibles también puede ser deducida de $S \cdot N$ por medio de las mismas hipótesis subsidiarias; y como el nuevo criterio de Ayer está formulado esencialmente en términos de la deducibilidad de un cierto tipo de consecuencias de la oración dada, apoya a $S \cdot N$ junto con S. Church ha señalado otra dificultad; ha mostrado,[9] en efecto, que si hay tres oraciones observacionales, ninguna de las cuales sola implica a ninguna de las otras, entonces se desprende, para cualquier oración S, que ella o su negación tiene contenido empírico según el criterio revisado de Ayer.

Todos los criterios considerados hasta ahora pretenden elucidar el concepto de significación empírica especificando ciertas conexiones lógicas que debe haber entre una oración significante y oraciones observacionales apropiadas. Parecería que este tipo de enfoque ofrece pocas esperanzas de lograr criterios precisos de significación: sugiere esta conclusión el anterior examen de algunos intentos representativos y recibe un refuerzo adicional de otras consideraciones, algunas de las cuales expondremos en las secciones que siguen.

3. *Caracterización de las oraciones significantes mediante criterios relativos a sus términos constituyentes*

Se presenta un procedimiento alternativo que, nuevamente, parece reflejar bien el punto de vista general del empirismo: sería posible caracterizar las oraciones cognoscitivamente significantes mediante ciertas condiciones que deban satisfacer sus componentes. Específicamente, parecería razonable decir que todos los términos extralógicos[10] de una oración sig-

[8] Esta restricción se expresa en forma recursiva y no contiene ningún círculo vicioso. Para la enunciación completa del criterio de Ayer, véase Ayer (1946), pág. 13.

[9] Church (1949). Un criterio alternativo recientemente propuesto por O'Connor (1950) como revisión de la formulación de Ayer está sujeto a una ligera variante de la crítica de Church: puede demostrarse que si hay tres oraciones observacionales, ninguna de las cuales implica a ninguna de las otras, y si S no es una oración compuesta, entonces S o $\sim S$ es significante de acuerdo con el criterio de O'Connor.

[10] Un término extralógico es aquel que no pertenece al vocabulario específico de la lógica. Las siguientes expresiones, y todas las definibles mediante ellas, son

nificante deben tener una referencia experiencial y que, por ende, sus significados deben ser susceptibles de elucidación con referencia a observables, de manera exclusiva.[11] Con el fin de mostrar ciertas analogías entre estos enfoques, adoptaremos las siguientes convenciones terminológicas.

A todo término que pueda aparecer en una oración cognoscitivamente significante lo llamaremos un *término cognoscitivamente significante*. Además, entenderemos por *término observacional* a todo término que (a) sea un *predicado observacional*, esto es, que significa alguna característica observable (como los términos "azul", "caliente", "suave", "coincidente con", "de mayor brillo aparente que", etc.), o (b) nombre algún objeto físico de tamaño macroscópico (como los términos "la aguja de este instrumento", "la Luna", "el volcán Krakatoa", "Greenwich, Inglaterra", "Julio César", etcétera).

Ahora bien, mientras que el criterio testabilista del significado pretendía caracterizar las oraciones cognoscitivamente significantes mediante ciertas conexiones inferenciales con algunas oraciones de observación, el enfoque alternativo que estamos considerando trata, en cambio, de especificar el vocabulario que puede usarse para formar oraciones significantes. Este vocabulario, la clase de los términos significantes, estaría caracterizado por la condición de que cada uno de sus elementos sea un término lógico o un término con significación empírica. En este último caso debe hallarse en ciertas relaciones definicionales o elucidativas con algunos términos observacionales. Este enfoque evita, ciertamente, toda violación de las anteriores condiciones de adecuación. Así, por ejemplo, si S es una oración significante, es decir, si contiene sólo términos cognoscitivamente significantes, entonces lo mismo sucederá con su negación, puesto que el signo de negación y sus equivalentes verbales pertenecen al vocabulario de la lógica y son, pues, significantes. De igual modo, si N es una oración que contiene un término no significante, entonces lo mismo sucederá con cualquier oración compuesta que contenga a N.

Pero esto no basta, por supuesto. Por el contrario, tendremos que considerar ahora un problema fundamental análogo al planteado por el enfoque anterior: ¿De qué manera precisa es menester concebir las conexiones lógicas entre los términos empíricamente significantes y los términos observacionales para obtener un criterio adecuado de significación cognoscitiva? Consideremos algunas posibilidades.

(3.1) El criterio más simple que se nos presenta podría ser llamado el *requisito de definibilidad*. Este exigiría que todo término con significación cognoscitiva deba ser definible explícitamente por medio de términos observacionales.

---

ejemplos típicos de términos lógicos: "no", "o", "si ... entonces", "todos", "algunos", " ... es un elemento de la clase ... ". Si es o no posible establecer una distinción teórica nítida entre términos lógicos y términos extralógicos es un controvertido problema que se relaciona con el de discriminar entre oraciones analíticas y sintéticas. Para nuestros fines, podemos suponer simplemente que el vocabulario lógico está dado por enumeración.

[11] Para una exposición y un examen crítico detallados de esta idea, véase el aclarador y estimulante artículo de H. Feigl (1950).

Este criterio parecería adecuarse al principio del operacionalismo según el cual todos los términos significantes de las ciencias empíricas deben ser introducidos por definiciones operacionales. Sin embargo, el requisito de definibilidad es demasiado restrictivo, pues no es posible definir explícitamente en términos observacionales muchos términos importantes del discurso científico y aun precientífico.

En efecto, como ha señalado Carnap,[12] el intento de suministrar definiciones explícitas en términos de observables halla serias dificultades tan pronto como es menester explicar términos disposicionales como "soluble", "maleable", "conductor eléctrico", etc. y muchos de ellos aparecen ya en el nivel del discurso precientífico.

Consideremos, por ejemplo, la palabra "frágil". Podríamos tratar de definirla diciendo que un objeto x es frágil si y sólo si satisface la siguiente condición: si en el momento t se golpea al objeto con fuerza, entonces se rompe en ese momento. Pero si se conciben como extensionales los conectivos de esta formulación, de modo que pueda simbolizarse la definición así:

$$\text{(D)} \qquad Fx \equiv (t)\ (Gxt \supset Rxt),$$

entonces el predicado "F" así definido no tiene el significado que se pretendía. En efecto, sea a cualquier objeto que no es frágil (por ejemplo, una gota de lluvia o una cinta de goma), pero que no es golpeado con fuerza en ningún momento de toda su existencia. Entonces "Gat" es falsa y, por ende, "$Gat \supset Rat$" es verdadera para todos los valores de "t"; por consiguiente, "Fa" es verdadera aunque a no es frágil.

Para remediar este defecto, podríamos interpretar la frase "si ... entonces..." del definiens original en un sentido más restrictivo que el condicional extensional. Podría sugerir este sentido la frase subjuntiva "si x fuera golpeado con fuerza en cualquier momento t, entonces x se rompería en t". Pero una elaboración satisfactoria de esta interpretación exigiría una clarificación del significado y la lógica de los condicionales contrafácticos y subjuntivos, lo cual constituye un espinoso problema.[13]

Carnap sugirió un procedimiento alternativo en su teoría de las oraciones de reducción.[14] Estas son oraciones que, a diferencia de las definiciones, especifican el significado de un término sólo condicional o parcialmente. El término "frágil", por ejemplo, podría ser introducido mediante la siguiente oración de reducción:

$$\text{(R)} \qquad (x)\ (t)\ [Gxt \supset (Fx \equiv Rxt)]$$

[12] Véase (1936-1937), en particular la sección 7.

[13] Sobre este tema, véase, por ejemplo, Langford (1941); Lewis (1946), págs. 210-230; Chisholm (1946); Goodman (1947); Reichenbach (1947), capítulo VIII; Hempel y Oppenheim (1948), parte III. Popper (1949); y en particular el análisis más detallado de Goodman (1955).

[14] Véase Carnap, *loc. cit.*, nota 11. Para una breve exposición elemental de la idea principal, véase Carnap (1938), parte III. La oración R formulada aquí para el predicado "F" sólo ejemplifica el tipo más simple de oración de reducción, la llamada oración de reducción bilateral.

que especifica que si x es golpeado con fuerza en cualquier momento t, entonces x es frágil si y sólo si se rompe en el momento t.

Ahora evitamos nuestra anterior dificultad, pues si a es un objeto no frágil que nunca es golpeado con fuerza, entonces esa expresión de R que sigue a los cuantificadores es verdadera de a; pero esto no implica que "*Fa*" sea verdadera. Sin embargo, la oración de reducción R sólo especifica el significado de "*F*" en su aplicación a aquellos objetos que satisfacen la "condición de testeo" de ser golpeados con fuerza en algún momento. Para los objetos que satisfacen esta condición, el significado de "*F*" queda indeterminado. En este sentido, las oraciones de reducción tienen el carácter de definiciones parciales o condicionales.

Las oraciones de reducción suministran una interpretación satisfactoria del contenido experiencial de una clase grande de términos disposicionales y permiten una formulación más adecuada de las llamadas definiciones operacionales, las cuales, en general, no son en modo alguno definiciones completas. Estas consideraciones sugieren una alternativa muy liberal del requisito de la definibilidad:

(3.2) *Requisito de reducibilidad.* Todo término con significación empírica debe poderse introducir, sobre la base de términos observacionales, a través de cadenas de oraciones de reducción.

Este requisito es característico de las versiones liberalizadas del positivismo y del fisicalismo que, desde 1936, aproximadamente, han superado la vieja concepción, demasiado estrecha, de la plena definibilidad de todos los términos de las ciencias empíricas mediante observables,[15] y evita muchos de los inconvenientes de ésta. Sin embargo, las oraciones de reducción no parecen ofrecer un medio adecuado para la introducción de los términos fundamentales de las teorías científicas avanzadas, a menudo llamadas construcciones teóricas. Revelan esto las siguientes consideraciones: una cadena de oraciones de reducción brinda una condición necesaria y suficiente para la aplicabilidad del término que introduce (cuando las dos condiciones coinciden, la cadena equivale a una definición explícita). Pero tomemos, por ejemplo, el concepto de longitud tal como se lo usa en la teoría física clásica. En ésta, la longitud en centímetros de la distancia entre dos puntos puede asumir como valor cualquier número real positivo; sin embargo, es claramente imposible formular por medio de términos observacionales una condición suficiente para la aplicabilidad de expresiones tales como "tener una longitud de $\sqrt{2}$ cm" o "tener una longitud de $\sqrt{2} + 10^{-100}$ cm", pues tales condiciones tendrían que brindar la posibilidad de discriminar, en términos observacionales, entre dos longitudes que difieren sólo en 10^{-100} cm.[16]

No sería aconsejable sostener que, por esta razón, sólo debemos admitir tales valores de la magnitud longitud que permitan la formulación de condiciones suficientes en términos de observables, pues esto excluiría,

[15] Véase el análisis de Carnap (1936-1937), en particular la sección 15; véase también la presentación más breve del punto de vista liberalizado en Carnap (1938).

[16] (Agregada en 1964.) Esto no es estrictamente correcto. Para una formulación más circunspecta véase la nota 13 del capítulo IV y el examen más completo de la sección 7 del capítulo VIII.

entre otras cosas, todos los números irracionales y nos impediría asignar a la diagonal de un cuadrado de lado igual a 1 la longitud $\sqrt{2}$, tal como lo exige la geometría euclidiana. En tal caso, los principios de la geometría euclidiana no serían aplicables en la física. Análogamente, serían inaplicables los principios del cálculo infinitesimal; y el sistema de la teoría científica tal como la conocemos hoy se reduciría a una construcción trunca, torpe e inútil. Por ello, no es así como puede resolverse la dificultad. En cambio, tendremos que analizar más detenidamente la función de las construcciones teóricas en las teorías científicas, con vista a obtener mediante tal análisis una caracterización más apropiada de los términos cognoscitivamente significantes.

Las construcciones teóricas aparecen en la formulación de las teorías científicas. Puede concebirse a éstas, en sus etapas avanzadas, como enunciadas en la forma de sistemas axiomáticos desarrollados deductivamente. La mecánica clásica y la interpretación física de la geometría euclidiana y algunos tipos de geometrías no euclidianas son ejemplos de tales sistemas. Los términos extralógicos usados en una teoría de este tipo pueden dividirse, a la manera habitual, en términos primitivos o básicos, no definidos dentro de la teoría, y términos definidos explícitamente mediante los términos primitivos. Así, por ejemplo, en la axiomatización de Hilbert de la geometría euclidiana, los términos "punto", "recta", "entre" se cuentan entre los primitivos, mientras que "segmento de recta", "ángulo", "triángulo" y "longitud" se cuentan entre los términos definidos. Los términos básicos y los definidos, juntamente con los términos de la lógica, constituyen el vocabulario con el cual se construyen todas las oraciones de la teoría. Estas últimas se dividen, en su presentación axiomática, en enunciados primitivos (también llamados postulados o enunciados básicos), que no derivan —dentro de la teoría— de otros enunciados, y enunciados derivados, que se obtienen por deducción lógica a partir de los primitivos.

Una teoría axiomatizada puede desarrollarse, a partir de los términos y oraciones primitivas, mediante principios puramente formales de definición y deducción, sin consideración alguna referente a la significación empírica de sus términos extralógicos. En realidad, éste es el procedimiento típico empleado en el desarrollo axiomático de teorías matemáticas no interpretadas como la de los grupos abstractos, anillos o látices, o cualquier tipo de geometría pura (esto es, no interpretada).

Sin embargo, un sistema desarrollado deductivamente de este tipo puede constituir una teoría científica sólo si ha recibido una interpretación empírica[17] que lo relacione con los fenómenos de nuestra experiencia. Se realiza tal interpretación asignando un significado, en términos de obser-

[17] La interpretación de las teorías formales ha sido intensamente estudiada por Reichenbach, particularmente en sus precursores análisis del espacio y el tiempo en la física clásica y en la relativista. Describe tal interpretación como el establecimiento de *definiciones coordinantes* (Zuordnungs definitionen) para ciertos términos de la teoría formal. Véase, por ejemplo, Reichenbach (1928). Más recientemente, Northrop [véase (1947), cap. VII, y también el detallado estudio del uso de teorías formuladas deductivamente en la ciencia, *ibíd.*, caps. IV, V y VI] y H. Margenau [véase, por ejemplo (1935)] han examinado ciertos aspectos de este proceso con el nombre de *correlación epistémica.*

vables, a ciertos términos u oraciones de la teoría formalizada. Con frecuencia no se da una interpretación a los términos y enunciados primitivos sino, más bien, a algunos de los términos definibles por medio de los primitivos, o algunas de las oraciones deducibles de los postulados.[18] Además, la interpretación puede ser solamente una asignación parcial de significado. Por ejemplo, puede considerarse que las reglas para la medición de longitudes por medio de una vara patrón suministran una interpretación empírica *parcial* del término "la longitud, en centímetros, del intervalo *i*" o, alternativamente, de algunas oraciones de la forma "la longitud del intervalo *i* es *r* centímetros". En efecto, el método sólo es aplicable a intervalos de cierto tamaño medio, y aun para éstos no constituye una interpretación completa, pues el uso de una vara patrón no es la única manera de determinar la longitud: hay varios procedimientos alternativos que suponen la medición de otras magnitudes vinculadas, mediante leyes generales, con la longitud que se quiere determinar.

Esta última observación concerniente a la posibilidad de una medición indirecta de la longitud en virtud de ciertas leyes sugiere una importante advertencia. No es correcto hablar, como se hace a menudo, del "significado experiencial" de un término o una oración aisladamente. En el lenguaje de la ciencia, y por razones similares aun en el discurso precientífico, un enunciado aislado por lo común no tiene implicaciones experienciales. En una teoría científica, una oración aislada, por lo común, no implica oraciones observacionales; sólo es posible deducir de ella consecuencias que afirman la producción de ciertos fenómenos observables uniéndola a un conjunto de otras hipótesis subsidiarias. De éstas, algunas serán habitualmente oraciones observacionales, otras serán enunciados teóricos previamente aceptados. Así, por ejemplo, la teoría relativista de la deflexión de la luz en el campo gravitacional del Sol implica aserciones acerca de fenómenos observables sólo si se la une a un *corpus* considerable de teorías astronómicas y ópticas, así como a un gran número de enunciados específicos acerca de los instrumentos utilizados en esas observaciones de eclipses solares que sirven para testear la hipótesis aludida.

Por consiguiente, la frase "el significado experiencial de la expresión E" es elíptica. Lo que "significa" una expresión dada con respecto a datos empíricos potenciales depende de dos factores, a saber:

I. *El armazón lingüístico L* al cual pertenece la expresión. Sus reglas determinan, en particular, qué oraciones —observacionales de otro tipo— pueden inferirse a partir de un enunciado o una clase de enunciados.

II. El contexto teórico en el que aparece la expresión, es decir, la clase de aquellos enunciados de *L* que se usan como hipótesis subsidiarias.

Así, la oración que formula la ley de la gravitación de Newton no tiene

[18] Una exposición un poco más detallada de este tipo de interpretación puede hallarse en Carnap (1939), § 24. Los artículos de Spence (1944) y MacCorquodale y Meehl (1948) brindan esclarecedores ejemplos del uso de construcciones teóricas en un campo ajeno a las ciencias físicas y de las dificultades con que se tropieza en el intento de analizar en detalle su función e interpretación.

en sí misma significado experiencial; pero cuando se la usa en un lenguaje cuyos componentes lógicos permiten el desarrollo del cálculo infinitesimal y cuando se la combina con un adecuado sistema de otras hipótesis —incluyendo oraciones que vinculen algunos de los términos teóricos con términos observacionales y, de este modo, den una interpretación parcial—, entonces tiene referencia a fenómenos observables de una gran variedad de campos. Consideraciones análogas se aplican al término "campo gravitacional", por ejemplo. Sólo se puede considerar que tiene significado experiencial dentro del contexto de una teoría, que debe ser al menos parcialmente interpretada; y el significado experiencial del término —expresado, digamos, en la forma de criterios operacionales para su aplicación— dependerá nuevamente del sistema teórico en cuestión y de las características lógicas del lenguaje dentro del cual se lo formula.

4. *La significación cognoscitiva como característica de los sistemas interpretados*

Las consideraciones anteriores apuntan a la conclusión de que no es posible lograr un criterio satisfactorio de significación cognoscitiva mediante el segundo enfoque de los considerados hasta ahora, es decir, mediante requisitos específicos para los términos que constituyan oraciones significantes. Este resultado concuerda con una característica general de la teorización científica (y aun, en principio, precientífica): la formación teórica y la formación conceptual marchan a la par; ninguna de ellas puede realizarse exitosamente aislada de la otra.

Por lo tanto, para poder atribuir significación cognoscitiva a algo, entonces ello sólo podrá ser con referencia a sistemas teóricos totales formulados en un lenguaje con una estructura bien determinada. Y la señal decisiva de la significación cognoscitiva en tales sistemas parece ser la interpretación de ellos en términos de observables. Podrá formularse tal interpretación, por ejemplo, mediante oraciones condicionales o bicondicionales que vinculen términos no observacionales del sistema con términos observacionales del lenguaje dado; estos últimos, así como las oraciones que establecen el vínculo, pueden o no pertenecer al sistema teórico.

Pero el requisito de la interpretación parcial es extremadamente liberal. Lo satisface, por ejemplo, el sistema consistente en la teoría física contemporánea combinado con un conjunto de principios de la metafísica especulativa, aunque ésta no tenga ninguna interpretación empírica. Dentro del sistema total, estos principios metafísicos desempeñan el papel de lo que K. Reach y también O. Neurath solían llamar *oraciones aisladas*: no son verdades o falsedades puramente formales, demostrables o refutables mediante las reglas lógicas del sistema lingüístico dado, ni tienen ningún contenido experiencial. Es decir, su omisión en el sistema teórico no tendría ningún efecto sobre su poder explicativo y predictivo con respecto a fenómenos potencialmente observables (esto es, el tipo de fenómenos que describen las oraciones observacionales). ¿No deberíamos, pues, exigir que un sistema cognoscitivamente significante no contenga oraciones aisladas? Se presenta el siguiente criterio:

(4.1) Un sistema teórico es cognoscitivamente significante si y sólo si está parcialmente interpretado, al menos en la medida en que ninguna de sus oraciones primitivas está aislada.

Pero este requisito puede excluir de un sistema teórico ciertas oraciones que bien pueden ser consideradas como admisibles y hasta deseables. A manera de simple ejemplo, supongamos que nuestro sistema teórico T contiene la oración primtiva:

$$(S1) \qquad (x)[P_1x \supset (Qx \equiv P_2x)]$$

donde "P_1" y "P_2" son predicados observacionales del lenguaje dado, L, mientras que "Q" funciona en T de manera algo similar a la de una construcción teórica y aparece sólo en una oración primitiva de T, a saber $S1$. Ahora bien, $S1$ no es una verdad o una falsedad de la lógica formal; y además, si se omite S_1 del conjunto de oraciones primitivas de T, entonces el sistema resultante, T', posee exactamente el mismo poder sistemático, esto es, explicativo y predictivo, que T. El criterio estudiado, pues, consideraría a $S1$ como una oración aislada que debe ser eliminada —cortada con la navaja de Occam, por así decirlo— para que el sistema teórico sea cognoscitivamente significante.

Pero es posible adoptar un punto de vista mucho más liberal de $S1$, considerándola como una definición parcial del término teórico "Q". Así concebida, $S1$ especifica que en todos los casos en que está presente la característica observable P_1, "Q" es aplicable si está presente también la característica observable P_2. En realidad, $S1$ es un caso de esas definiciones parciales o condicionales que Carnap llama oraciones de reducción bilaterales. Carnap califica explícitamente estas oraciones de analíticas (aunque no, por supuesto, de verdades de la lógica formal), esencialmente por la razón de que todas sus consecuencias expresables mediante predicados observacionales (y términos lógicos) solamente son verdades de la lógica formal.[19]

Desarrollemos un poco más esta línea de pensamiento. Esto nos llevará a algunas observaciones sobre las oraciones analíticas y luego otra vez a la cuestión de la adecuación de (4.1).

Supongamos que agregamos a nuestro sistema T la siguiente oración:

$$(S2) \qquad (x)[P_3x \supset (Qx \equiv P_4x)]$$

donde "P_3" y "P_4" son predicados de observación adicionales. Entonces, según la tesis de que "toda oración de reducción bilateral es analítica",[20] $S2$ sería tan analítica como $S1$. Sin embargo, las dos oraciones implican conjuntamente consecuencias no analíticas que son expresables en términos de predicados de observación en forma exclusiva, tales como:[21]

$$(O) \qquad (x)[\sim (P_1x \cdot P_2x \cdot P_3x \cdot \sim P_4x) \cdot \sim \sim (P_1x \cdot \sim P_2x \cdot P_3x \cdot P_4x)]$$

19 Véase Carnap (1936-1937), en particular las secciones 8 y 10.

20 Carnap (1936-1937), pág. 452.

21 La oración O es lo que Carnap llama la *oración representativa* del par consistente en las oraciones $S1$ y $S2$; véase (1936-1937), págs. 450-453.

Pero no podemos admitir la consecuencia de que la conjunción de dos oraciones analíticas pueda ser sintética. Luego, para que el concepto de analiticidad pueda aplicarse a las oraciones de sistemas deductivos interpretados será menester relativizarlo con respecto al contexto teórico. Así, por ejemplo, *S*1 podría ser calificada de analítica con respecto al sistema *T*, cuyos postulados restantes no contienen el término "*Q*", pero de sintética con respecto al sistema *T* enriquecido con *S*2. Hablando en términos estrictos, es necesario relativizar el concepto de analiticidad también con respecto a las reglas del lenguaje usado, pues éstas determinan las consecuencias observacionales o de otro tipo implicadas por una sentencia dada. Esta necesidad de realizar al menos una doble relativización del concepto de analiticidad era casi de esperar en vista de las consideraciones que exigían la misma doble relativización del concepto de significado experiencial de una oración.

Por otra parte, si decidimos no admitir *S*1 en el papel de una definición parcial y, en lugar de esto, la rechazamos por considerarla una oración aislada, entonces llegamos a una conclusión análoga que una oración sea o no aislada depende del marco lingüístico y del contexto teórico adoptados; si bien *S*1 es aislada con respecto a *T* (y el lenguaje en el que ambas se formulan), adquiere implicaciones experienciales definidas cuando se amplía T con *S*2.

Encontramos, así, en el nivel de los sistemas teóricos interpretados, una peculiar semejanza y una fusión parcial de algunos de los problemas atinentes a los conceptos de significación cognoscitiva y de analiticidad: ambos conceptos deben ser relativizados, y una clase grande de oraciones puede ser considerada, en apariencia con igual derecho, como analítica en un contexto dado o como aislada, o no significante, con respecto a él.

Además de suprimir, por considerarlas aisladas en un contexto dado, ciertas oraciones que podrían ser concebidas como definiciones parciales, el criterio (4.1) tiene otro defecto serio. De dos formulaciones lógicamente equivalentes de un sistema teórico, puede calificar a uno de significante y excluir al otro por contener una oración aislada entre sus oraciones primitivas. En efecto, supongamos que un cierto sistema teórico *T*1 contiene entre sus oraciones primitivas *S′*, *S″*, ... exactamente una, *S′*, que está aislada. Entonces *T*1 no es significante, según (4.1). Pero consideremos ahora el sistema teórico *T*2 obtenido de *T*1 y al reemplazar las dos primeras oraciones primitivas *S′* y *S″* por una sola, a saber, su conjunción. Entonces, según nuestras suposiciones, ninguna de las oraciones primitivas de *T*2 está aislada, y *T*2, aunque equivalente a *T*1, es calificada de significante por (4.1). Con el fin de hacer justicia a la intención de (4.1), pues, debemos establecer el siguiente requisito, más estricto:

(4.2) Un sistema teórico es cognoscitivamente significante si y sólo si es interpretado de manera parcial en una medida tal que en ningún sistema equivalente a él al menos una oración primitiva está aislada.

Apliquemos este requisito a un sistema teórico cuyos postulados incluyen las dos oraciones *S*1 y *S*2 consideradas antes, y cuyos otros postulados no contengan "*Q*". Puesto que las oraciones *S*1 y *S*2 implican en

conjunto la oración *O*, el conjunto consistente en *S*1 y *S*2 es lógicamente equivalente al conjunto formado por *S*1, *S*2 y *O*. Luego, si reemplazamos el primer conjunto por el último, obtenemos un sistema teórico equivalente al dado. En este nuevo sistema, tanto *S*1 como *S*2 están aisladas, ya que puede demostrarse que su eliminación no afecta al poder explicativo y predictivo del sistema con referencia a fenómenos observables. Para decirlo en términos intuitivos, el poder sistemático de *S*1 y *S*2 es igual al de *O*. Por ende, el sistema original queda descalificado por (4.2). Desde el punto de vista de un positivismo estrictamente sensorialista, como el concebido quizá por Mach, puede darse la bienvenida a este resultado como un sano repudio de las teorías que hacen referencia a entidades ficticias y una estricta adhesión a teorías expresadas exclusivamente en términos de observables. Pero desde un punto de vista moderno, debemos decir que tal procedimiento pasa por alto o juzga erróneamente la importante función de las construcciones teóricas en la teoría científica: la historia de la empresa científica demuestra que si queremos llegar a leyes generales, precisas, amplias y bien confirmadas, tenemos que elevarnos por encima del nivel de la observación directa. Los fenómenos directamente accesibles a nuestra experiencia no están vinculados por leyes generales de gran alcance y rigor. Para la formulación de tales leyes de nivel superior se necesitan las construcciones teóricas. Una de las funciones más importantes de una construcción teórica bien elegida es su capacidad potencial de servir como componente de las nuevas conexiones generales que puedan descubrirse. Y estaríamos ciegos para tales conexiones si insistiéramos en desterrar de las teorías científicas todos aquellos términos y oraciones de los que puede "prescindirse" en el sentido indicado en (4.2). Si adoptáramos tal estrecha posición fenomenalista o positivista, renunciaríamos a la enorme fertilidad de las construcciones teóricas y a menudo daríamos a la estructura formal de la teoría expurgada un carácter engorroso e ineficaz.

Debe abandonarse, pues, el criterio (4.2), y consideraciones como las esbozadas en este capítulo parecen prestar fuerte apoyo a la conjetura de que no puede hallarse ninguna adecuada alternativa del mismo; es decir, que no es posible formular criterios generales y precisos que permitan separar los sistemas parcialmente interpretados de cuyas oraciones aisladas podría decirse que tienen una función significante de aquellos en los que las oraciones aisladas son, por así decirlo, meros agregados inútiles.

Concluimos antes que la significación cognoscitiva, en el sentido que dan a la expresión el empirismo y el operacionalismo actuales, a lo sumo puede atribuirse a oraciones que formen un sistema teórico, y quizá más bien a tales sistemas como un todo. Ahora bien, en lugar de tratar de reemplazar (4.2) por alguna alternativa, tendremos que reconocer que la significación cognoscitiva, en un sistema, es una cuestión de grados: los sistemas significantes van desde aquellos cuyo vocabulario extralógico total consiste en términos de observación, pasando por teorías cuya formulación reposa en gran medida en construcciones teóricas, hasta sistemas que apenas tienen alguna vinculación con hallazgos empíricos potenciales. En lugar de dicotomizar este conjunto en sistemas significantes y sistemas no significantes parece menos arbitrario y más promisorio evaluar o comparar

diferentes sistemas teóricos en lo que respecta a características tales como las siguientes:

a) La claridad y precisión con que se formulan las teorías y con que se han hecho explícitas las relaciones lógicas entre sus elementos y con expresiones dadas en términos observacionales.
b) El poder sistemático —es decir, explicativo y predictivo— de los sistemas con respecto a fenómenos observables.
c) La simplicidad formal del sistema teórico con el que se logra cierto poder sistemático.
d) La medida en que las teorías han sido confirmadas por elementos de juicio experienciales.

Muchos de los enfoques filosóficos especulativos de la cosmología, la biología o la historia, por ejemplo, harían un pobre papel prácticamente en todos esos ítems, y luego revelarían su incapacidad para sostener la comparación con las teorías rivales disponibles o demostrarían ser tan poco promisorios que no justificarían su ulterior estudio o desarrollo.

Si se desarrolla en detalle el procedimiento aquí sugerido, de modo que se haga aplicable también a casos menos obvios, entonces será necesario, por supuesto, elaborar normas generales y teorías correspondientes a ellas, para la evaluación y comparación de sistemas teóricos en los diversos aspectos mencionados. No es posible juzgar de antemano en qué medida puede hacerse esto con rigor y precisión. En años recientes es mucho lo que se ha trabajado para llegar a una definición y una teoría del concepto de grado de confirmación, o probabilidad lógica, de un sistema teórico,[22] y se han hecho varias contribuciones tendientes a la clarificación de algunas de las ideas a las que nos hemos referido.[23] La continuación de esta labor investigativa es un estímulo para realizar un mayor trabajo constructivo en el análisis lógico y metodológico del conocimiento científico.

POSTSCRIPTUM (1964) SOBRE LA SIGNIFICACION COGNOSCITIVA

El capítulo precedente es una fusión de dos artículos: "Problems and Changes in the Empiricist Criterion of Meaning", *Revue Internationale de Philosophie* Nº 11 (1950), y "The Concept of Cognitive Significance: A Reconsideration", *Proceedings of the American Academy of Arts and Sciences* 80 (1951). Al combinar ambos artículos, he omitido sobre todo algunas partes del primero que habían quedado ampliamente superadas por el segundo;[24] hice también algunos cambios menores en el texto res-

[22] Véase, por ejemplo, Carnap (1945)1 y (1945)2, y sobre todo (1950). Véase también Helmer y Oppenheim (1945).

[23] Sobre la simplicidad, consultar en particular Popper (1935), cap. V; Reichenbach (1938), § 42; Goodman (1949), (1949)2 y (1950); sobre el poder explicativo y predictivo, véase Hempel y Oppenheim (1948), parte IV.

[24] Las ideas básicas presentadas en los artículos anteriores y en esta versión refundida de ellos han sido sometidas a un penetrante examen por I. Scheffler en *The Anatomy of Inquiry*, Nueva York, 1963.

tante. Algunos de los problemas generales planteados en el ensayo combinado se desarrollan aún más en otros trabajos, en particular en "El dilema del teórico", en el capítulo VIII de esta obra. En este postcriptum me propongo simplemente señalar algunas ideas secundarias concernientes a puntos particulares del ensayo anterior.

(I) Las objeciones 2.1(c) y 2.2(c) contra los requisitos de verificabilidad completa y refutabilidad completa son, creo, de fuerza discutible. En efecto, sólo puede decirse propiamente que O implica $O \vee N$, y $O \cdot N$ a su vez implica O, si tanto N como O son oraciones asertivas y, por lo tanto, verdaderas o falsas. Pero si se entiende el criterio de significación cognoscitiva en el sentido de que delimita la clase de las oraciones que hacen aserciones significantes y que son, por ende, verdaderas o falsas, entonces la oración N aducida en las objeciones no es asertiva, ni lo son $S \vee N$ o $S \cdot N$; luego, las presuntas inferencias de $S \cdot N$ a S y de S a $S \vee N$ son inadmisibles.[25]

Mi objeción mantiene su fuerza, en cambio, contra el uso de la refutabilidad, no como criterio de significación, sino como "criterio de demarcación". Este uso trazaría una línea divisoria "entre los enunciados o sistemas de enunciados de las ciencias empíricas y todos los otros enunciados, sean de carácter religioso, metafísico o simplemente seudocientífico".[26] En efecto, el argumento 2.2(c) muestra que la conjunción de un enunciado científico S con un enunciado no científico N es refutable y, por consiguiente, lo califica como enunciado científico; y esto frustra el propósito al que apunta el criterio de demarcación.

(II) Mi aserción, en 2.1(a) y 2.2(a), de que los requisitos de verificabilidad y de refutabilidad excluirían *todas* las hipótesis de forma cuantificacional mixta es falsa. Consideremos la hipótesis "todos los cuervos son negros y algo es blanco", o, en símbolos:

$$(x)(Cx \supset Nx) \cdot (\exists y) By,$$

que es equivalente a

$$(x)(\exists y)[(Cx \supset Nx) \cdot By]$$

Esta oración satisface el requisito de refutabilidad porque implica la hipótesis universal pura "$(x)(Cx \supset Nx)$", que sería refutada, por ejemplo, por el siguiente conjunto de oraciones observacionales: ["Ca", "$\sim Na$"]. Análogamente, la oración:

$$(\exists x)(y)(Cx \vee By)$$

es verificable, pues es implicada, por ejemplo, por "Ca".

[25] Debo esta corrección a estudiantes egresados que expusieron la crítica anterior en uno de mis seminarios. El mismo punto ha sido formulado recientemente con mucha claridad por D. Rynin en "Vindication of L*G*C*LP*S*T*V*SM", *Proceedings and Addresses of the American Philosophical Association*, 30 (1957). Véase, en particular, las págs. 57-58.

[26] K. R. Popper, "Philosophy of Science. A Personal Report", en C. A. Mace (comp.), *British Philosophy in the Mid-Century*, Londres, 1957, págs. 155-191; las citas son de las págs. 163 y 162.

El punto esencial de la objeción, sin embargo, queda en pie: muchas hipótesis científicas de forma cuantificacional mixta no son verificables ni refutables, por lo que quedarían descalificadas tanto por el requisito de verificabilidad como por el de refutabilidad; y si se usa este último como criterio de demarcación y no de significación, excluye a esas hipótesis de la clase de los enunciados científicos. Estas consecuencias son inaceptables.

(III) Una crítica aún más enérgica de los criterios de verificabilidad y de refutabilidad resulta de la condición (A1) enunciada en la sección 2, que exige, en efecto, de todo criterio aceptable de significación que, si admite como significante una oración, admita también su negación como tal. Es obvia la necesidad de satisfacer esta condición, pues siendo una oración significante una oración que debe ser verdadera o falsa, no puede sostenerse que su negación no es significante sin violar un principio fundamental de la lógica. Aunque el criterio de refutabilidad sea usado como criterio de demarcación, y no de significación cognoscitiva, parece imperativa la satisfacción de la condición (A1). De otro modo, un científico que informara haber tenido éxito en la refutación de una hipótesis científica S de forma universal haría un enunciado no científico al decir: "luego, no se da el caso de que S sea válida", pues este enunciado no sería refutable. Con mayor generalidad, la inferencia lógica deductiva formalmente válida conduciría con frecuencia de premisas científicas a conclusiones no científicas —por ejemplo, de "$Ca \cdot \sim Na$" a "$(\exists x)(Cx \cdot \sim Nx)$"— lo cual, sin duda, es intolerable.

Pero cuando se combina el requisito de verificabilidad, o el de refutabilidad, con la condición (A1), entonces una oración es cognoscitivamente significante en el caso de que tanto ella como su negación sean verificables, o en el caso de que ella y su negación sean refutables. Los dos criterios exigen ahora lo mismo de una oración significante, a saber, que sea verificable y refutable. Esta caracterización admite, además de todos los compuestos extensionales de oraciones observacionales, también ciertas oraciones que contienen cuantificadores. Por ejemplo, "$Pa \vee (x)Qx$" es verificable por "Pa" y refutable por ["$\sim Pa$", "$\sim Qb$"]; y como puede verse fácilmente, "$Pa \cdot (\exists x)Qx$" cumple también con el requisito combinado. Pero este requisito excluye todas las hipótesis estrictamente generales, es decir, aquellas en las que los cuantificadores figuran esencialmente, pero no constantes de individuos, tales como "$(x)(Cx \supset Nx)$", "$(x)(\exists y)(Cxy \supset Sxy)$", etc. Sin duda, esta consecuencia también es inaceptable, independientemente de que se use el criterio para delimitar la clase de las oraciones significantes o la de los enunciados de las ciencias empíricas.

BIBLIOGRAFIA

Ayer, A. J.: *Language, Truth and Logic.* Londres, 1936; 2ª ed. 1946.

Carnap, R.: "Testability and Meaning." *Philosophy of Science,* 3 (1936) y 4 (1937).

Carnap, R.: "Logical Foundations of the Unity of Science." *International Enciclopedia of Unified Science,* I, 1. Chicago, 1938.

Carnap, R.: *Foundations of Logic and Mathematics.* Chicago, 1939.

Carnap, R.: "On Inductive Logic." *Philosophy of Science,* 12 (1945). Citado como (1945)1 en este trabajo.

Carnap, R.: "The Two Concepts of Probability." *Philosophy and Phenomenological Research,* 5 (1945). Citado como (1945)2 en este trabajo.

Carnap, R.: *Logical Foundations of Probability.* Chicago, 1950.

Chisholm, R. M.: "The Contrary-to-Fact Conditional". *Mind,* 55 (1946).

Church, A.: "Review of Ayer (1946)". *The Journal of Symbolic Logic,* 14 (1949), págs. 52-53.

Feigl, H.: "Existential Hypotheses: Realistic vs. Phenomenalistic Interpretations". *Philosophy of Science,* 17 (1950).

Goodman, N.: "The Problem of Counterfactual Conditionals". *The Journal of Philosophy,* 44 (1947).

Goodman, N.: "The Logical Simplicity of Predicates". *The Journal of Symbolic Logic,* 14 (1949). Citado como (1949)1 en este trabajo.

Goodman, N.: "Some Reflections on the Theory of Systems". *Philosophy and Phenomenological Research,* 9 (1949). Citado como (1949)2 en este trabajo.

Goodman, N.: "An Improvement in the Theory of Simplicity". *The Journal of Symbolic Logic,* 15 (1950).

Goodman, N.: *Fact, Fiction, and Forecast.* Cambridge, Massachusetts, 1955.

Helmer, O. y P. Oppenheim: "A Syntactical Definition of Probability and of Degree of Confirmation". *The Journal of Symbolic Logic,* 10 (1945).

Hempel, C. G. y P. Oppenheim: "Studies in the Logic of Explanation". *Philosophy of Science,* 15 (1948).

Langford, C. H.: Análisis publicado en *The Journal of Symbolic Logic,* 6 (1941), págs. 67-68.

Lewis, C. I.: *An Analysis of Knowledge and Valuation.* La Salle, Ill., 1946.

MacCorquodale, K. y P. E. Meehl: "On a Distinction Between Hypothetical Constructs and Intervening Variables". *Psychological Review,* 55 (1948).

Margenau, H.: "Methodology of Modern Physics". *Philosophy of Science,* 2 (1935).

Northrop, F. S. C.: *The Logic of the Sciences and the Humanities,* Nueva York, 1947.

O'Connor, D. J.: "Some Consequences of Professor A. J. Ayer's Verification Principle". *Analysis,* 10 (1950).

Pap, A.: *Elements of Analytic Philosophy.* Nueva York, 1949.

Popper, K.: *Logik der Forschung.* Viena, 1935.

Popper, K.: "A Note on Natural Laws and So-Called 'Contrary-to-Fact Conditionals'". *Mind,* 58 (1949).

Reichenbach, H.: *Philosophie der Raum-Zeit-Lehre.* Berlín, 1928.

Reichenbach, H.: *Elements of Symbolic Logic.* Nueva York, 1947.

Russell, B.: *Human Knowledge,* Nueva York, 1948.

Schlick, M.: "Meaning and Verification". *Philosophical Review,* 45 (1936). Reimpreso en Feigl, H. y W. Sellars (comps.); *Readings in Philosophical Analysis.* Nueva York, 1949.

Spence, Kenneth W.: "The Nature of Theory Construction in Contemporary Psychology". *Psychological Review,* 51 (1944).

Capítulo V

Evaluación lógica del operacionalismo [1]

1. *Tesis básicas del operacionalismo*

El operacionalismo está estrechamente emparentado con el empirismo lógico en sus tesis fundamentales. Ambas escuelas de pensamiento han dado mucho énfasis al significado o contenido experiencial definido como condición necesaria del discurso objetivamente significante, y ambas han dedicado grandes esfuerzos a establecer criterios explícitos de significación experiencial. Pero el empirismo lógico ha considerado el contenido experiencial como una característica de enunciados —a saber, como su capacidad para someterse a testeo por experimento u observación—, mientras que el operacionalismo ha tendido a concebir el significado experiencial como una característica de conceptos o de los términos que los representan, a saber, como su capacidad para encuadrarse en definiciones operacionales.

Las ideas básicas del análisis operacional. Se concibe una definición operacional de un término como una regla según la cual el término debe aplicarse a un caso particular si la realización de operaciones específicas para este caso da determinado resultado característico. Por ejemplo. el término "más duro que" podría definirse operacionalmente por la regla según la cual se dirá que un trozo de mineral x es más duro que otro trozo, y, si la operación de trazar con x una raya sobre la superficie de y da como resultado una raya sobre ésta. Análogamente, se conciben los diferentes valores numéricos de una magnitud tal como la longitud, como definibles operacionalmente con referencia a los resultados de operaciones de medición específica. Para salvaguardar la objetividad de la ciencia, se exige que todas las operaciones mencionadas en este tipo de definición sean intersubjetivas, en el sentido de que observadores diferentes deben poder realizar "la misma operación" con razonable acuerdo en sus resultados.[2]

[1] Este capítulo es una versión ligeramente modificada de un artículo con el mismo título que apareció en *Scientific Monthly*, 79, págs. 215-220 (1954). Se lo reimprime aquí con la amable autorización del director de *Science*.

[2] P. W. Bridgman, "Some General Principles of Operational Analysis" y "Rejoinders and Second Thoughts", *Psychological Review*, 52, 246 y 281 (1945); "The Nature of Some of our Physical Concepts", *British Journal for the Philosophy of Science*, 1, 258 (1951).

P. W. Bridgman, el creador del análisis operacional, distingue varios tipos de operaciones que pueden invocarse al especificar los significados de términos científicos.[3] Las principales son: (I) las que él llama *operaciones instrumentales*, que consisten en el uso de aparatos diversos para la observación y la medición; (II) las operaciones con papel y lápiz, las operaciones verbales, los experimentos mentales, etc.; este grupo incluye, entre otras cosas, las técnicas de la inferencia matemática y lógica, y los experimentos imaginarios. Para mayor brevedad, pero también para sugerir una semejanza fundamental entre los procedimientos del segundo tipo, las llamaré *operaciones simbólicas*.

Los conceptos de operación y de definición operacional sirven para enunciar los principios básicos del análisis operacional, de los cuales tienen especial importancia los siguientes:

1) "Los significados son operacionales". Para comprender el significado de un término, debemos conocer los criterios operacionales de su aplicación,[4] y por ende todo término científico significativo debe permitir una definición operacional. Tal definición puede hacer referencia a ciertas operaciones simbólicas y debe siempre, en última instancia, hacer referencia a alguna operación instrumental.[5]

2) Para evitar ambigüedades, debe definirse todo término científico por medio de un único criterio operacional. Aunque dos procedimientos operacionales diferentes (por ejemplo, la manera óptica y la táctil de medir longitudes) den los mismos resultados, se los debe considerar como conceptos definitorios diferentes (por ejemplo, la longitud óptica y la longitud táctil), y se los debe distinguir terminológicamente porque la presunta coincidencia de los resultados se infiere a partir de datos experimentales, y es "inseguro" olvidar que nuevos y quizá más precisos datos experimentales pueden demostrar que la presunción es falsa.[6]

3) La insistencia en que los términos científicos tengan significados operacionales especificables sin ambigüedades sirve para asegurar la posibilidad de un test objetivo para las hipótesis formuladas mediante esos términos.[7] Se rechazan por considerárselas carentes de significado las hipótesis que no admiten un test operacional o, más bien, las preguntas que suponen formulaciones no testables: "Si una pregunta específica tiene significado, debe ser posible hallar operaciones mediante las cuales se le pueda dar respuesta. Se encontrará en muchos casos que no pueden existir tales operaciones, y la pregunta, por lo tanto, no tiene significado".[8]

El énfasis en el "significado operacional" en el discurso científicamente significativo ha brindado, sin duda, una crítica saludable de ciertos tipos de procedimientos de la filosofía y de las ciencias empíricas, y ha

[3] P. W. Bridgman: "Operational Analysis", *Philosophy of Science*, 5, 123 (1938); *British Journal for the Philosophy of Science*, 1, 258 (1951).

[4] P. W. Bridgman: *Philosophy of Science*, 5, 116 (1938).

[5] P. W. Bridgman: *British Journal for Philosophy of Science*, 1, 260 (1951).

[6] P. W. Bridgman: *The Logic of Modern Physics*, Nueva York, Macmillan, 1927, págs. 6 y 23-24. *Philosophy of Science*, 5, 121 (1938); *Psychological Review*, 52, 247 (1945); "The Operational Aspect of Meaning", *Synthèse*, 8, 255 (1950-1951).

[7] P. W. Bridgman, *Psychological Review*, 52, 246 (1945).

[8] P. W. Bridgman: *The Logic of Modern Physics*, pág. 28.

dado un vigoroso estímulo al pensamiento metodológico. Sin embargo, las ideas centrales del análisis operacional, tales como las han formulado sus defensores, son tan vagas que no constituyen una teoría concerniente a la naturaleza de los conceptos científicos, sino más bien un programa para la elaboración de tal teoría. Comparten esta característica con la exigencia del empirismo lógico de que todos los enunciados científicos significativos tengan contenido experiencial, de que éste consiste en la testabilidad por datos apropiados de observación directa y de que aquellas oraciones que no admiten ningún test deben ser descartadas como "seudohipótesis" sin significado. Tampoco estas ideas constituyen tanto una tesis o una teoría como un programa para una teoría que es necesario formular y desarrollar en términos precisos.

El intento de elaborar una teoría operacionalista de los conceptos científicos deberá abordar al menos dos problemas principales: el de dar una elucidación más precisa del concepto de definición operacional y el de si la definición operacional, en el sentido explicado, es realmente necesaria y apropiada para la introducción de todos los términos no observacionales de las ciencias empíricas.

Quiero presentar aquí, en un resumido esbozo, algunas consideraciones que se relacionan con estos problemas. Limitaremos el examen al vocabulario descriptivo, o extralógico, de las ciencias empíricas, y no trataremos por ende de las ideas de Bridgman concernientes al carácter de la lógica y la matemática.

2. *Una concepción más amplia de la definición operacional y del programa del análisis operacional*

Los términos "significado operacional" y "definición operacional", así como muchas de las afirmaciones que se hacen en escritos operacionalistas, sugieren que los criterios de aplicación de todo término científico deben referirse en última instancia al resultado de cierto tipo específico de manipulación del tema que se investiga. Evidentemente, esto sería demasiado restrictivo. Una definición operacional da significado experiencial al término que introduce porque nos permite decidir acerca de la aplicabilidad de dicho término a un caso dado observando la respuesta del caso en condiciones especificables de test. Tiene gran importancia para la práctica de la investigación científica saber si es posible producir a voluntad esas condiciones mediante "operaciones instrumentales" o si debemos esperar su aparición; pero ello no es esencial para asegurar un contenido experimental al término definido; lo que interesa para este último propósito es simplemente que las condiciones de test pertinentes y la respuesta requerida sean de tal tipo que investigadores diferentes puedan discernir, por observación directa y con un grado razonable de acuerdo, si en un caso dado se realizan las condiciones del test y si se obtiene la respuesta característica.

Así, una definición operacional del tipo más simple —una definición que, hablando en términos aproximados, sólo se refiera a las operaciones instrumentales— deberá ser concebida más ampliamente, como introdu-

ciendo un término mediante la estipulación de que se aplicará a todos los casos, y sólo a ellos, que en condiciones observables específicas, *S*, presenten una característica respuesta observable, *R*.

Sin embargo, no puede concebirse una definición operacional como especificando que el término aludido se aplica a un caso dado sólo si *S* y *R* realmente se producen en este caso. Por ejemplo, se afirma que los cuerpos físicos tienen masa, temperatura, carga, etc., aun en momentos en que no se miden estas magnitudes. Luego, una definición operacional de un concepto deberá ser entendida como que atribuye el concepto a todos aquellos casos que *presentarían* la respuesta característica si se *realizaran* las condiciones del test. Un concepto caracterizado de este modo, obviamente, no es "sinónimo del correspondiente conjunto de operaciones".[9] No constituye una característica manifiesta, sino potencial, es decir, una disposición a presentar determinada respuesta característica en condiciones de test especificadas.

Pero atribuir una disposición de este género a un caso en que la condición específica del test no se realiza (por ejemplo, atribuir la solubilidad en el agua a un terrón de azúcar que no se coloca realmente en agua) es hacer una generalización, lo cual supone un riesgo inductivo.

Así, la aplicación de un término definido operacionalmente a un caso del tipo considerado tendría que ser juzgado como "insegura", precisamente en el mismo sentido en que Bridgman insiste en que es "inseguro" suponer que dos procedimientos de medición que han dado los mismos resultados en el pasado continuarán dándolos en el futuro. Está claro, pues, que si debemos rechazar todo procedimiento que suponga un riesgo inductivo, no sólo no podríamos usar más de un criterio operacional para introducir un término dado, sino que tampoco podríamos aplicar nunca un término disposicional a un caso en el que no se realicen las condiciones manifiestas características de aplicación. Por consiguiente, en efecto, estaría prohibido el uso de conceptos disposicionales.

Podríamos agregar aquí unas pocas observaciones acerca de las operaciones no instrumentales aprobadas para la introducción, en especial, de términos teóricos. En los escritos operacionalistas, esos procedimientos simbólicos han sido caracterizados tan vagamente que permiten la introducción, mediante una adecuada elección de operaciones "verbales" o "mentales", de prácticamente todas las ideas que el análisis operacional debía prohibir por considerarlas desprovistas de significado. Para resolver esta dificultad, Bridgman ha sugerido una distinción entre operaciones "buenas" y "malas";[10] pero no ha suministrado un criterio claro para

[9] P. W. Bridgman, *ibíd.*, pág. 5. Posteriormente, Bridgman introdujo una reserva en su respuesta (*Philosophy of Science*, 5, 117 [1938]) a R. B. Lindsay, "A Critique of Operationalism in Physics" (*Philosophy of Science*, 4 [1937]). La reserva era esencial, sobre la base —muy diferente de la presentada en este capítulo— de que el significado operacional sólo es una característica necesaria, pero presumiblemente no suficiente, de los conceptos científicos.

[10] P. W. Bridgman, *Philosophy of Science*, 5, 126 (1938); "Some Implications of Recent Points of View in Physics", *Revue Internationale de Philosophie*, 3, pág. 484 (1949). La distinción que se quiere establecer entre operaciones buenas y malas queda oscurecida por el hecho de que, en el examen de Bridgman, el significado de "operación buena" varía de lo que podría describirse como una "operación cuyo uso

establecer esta distinción. Por consiguiente, esta idea no logra tapar el agujero en el dique operacionalista.

Para que los principios del operacionalismo permitan las construcciones teóricas de la ciencia pero excluyan otros tipos determinados de términos como carentes de significado experiencial u operacional, es necesario que el vago requisito de la definibilidad con referencia a operaciones simbólicas instrumentales y "buenas" sea reemplazado por una caracterización precisa de los tipos de oraciones que pueden usarse para introducir o especificar los sentidos de términos no observacionales "significativos" sobre la base del vocabulario observacional de la ciencia. Tal caracterización eliminaría la noción psicologista de operación mental en favor de una especificación de los conceptos y procedimientos lógico-matemáticos que serán admitidos en el contexto de la definición operacional.

La referencia que acabo de hacer al vocabulario observacional de la ciencia es esencial para la idea de definición operacional, pues las condiciones de test y la respuesta característica especificada en una definición operacional se describen en los términos de este vocabulario, por medio del cual, por lo tanto, se caracterizan en última instancia los significados de los términos definidos operacionalmente. Por consiguiente, se respetará la intención de la insistencia operacionalista original en el carácter repetible intersubjetivo de las operaciones definitorias si exigimos que los términos incluidos en el vocabulario observacional se refieran a características directa y públicamente observables, esto es, cuya presencia o ausencia puede establecerse, en condiciones adecuadas, por observación directa y con un satisfactorio acuerdo entre observadores diferentes.[11]

En suma, pues, una enunciación y elaboración precisa de las afirmaciones básicas del operacionalismo exige una elucidación de las relaciones lógicas entre los términos teóricos y los observacionales, así como una enunciación y elaboración precisa de las afirmaciones básicas del empirismo exige una elucidación de las relaciones lógicas que vinculan las oraciones teóricas con las oraciones observacionales que describen datos potenciales de observación directa.

3. *Especificación del significado por definición explícita y por reducción*

En un comienzo puede parecer plausible suponer que todos los términos teóricos usados en la ciencia pueden ser definidos totalmente por

---

en la definición operacional asegura un significado y una testabilidad experiencial" a "procedimientos científicos —en algún sentido muy amplio— que nos conduce a hacer predicciones correctas".

[11] La condición así impuesta al vocabulario observacional de la ciencia sólo tiene un carácter pragmático: exige que cada término incluido en este vocabulario sea de tal tipo que, en condiciones apropiadas, diferentes observadores puedan llegar, mediante la observación directa, a un alto grado de acuerdo en lo referente a si el término se aplica a una situación determinada. La expresión "coincide con" aplicable a las agujas y a las marcas de las escalas de los aparatos, es un ejemplo de un término que satisface esta condición. El hecho de que los seres humanos sean capaces de elaborar vocabularios observacionales que satisfagan el requisito dado es una circunstancia afortunada: sin ella, sería imposible la ciencia como empresa intersubjetiva.

medio del vocabulario observacional. Hay varias razones, sin embargo, para dudar de esta suposición.

Ante todo, existe una dificultad concerniente a la definición de los términos científicos que se refieren a disposiciones, y como se ha señalado en un párrafo anterior, todos los términos introducidos por definiciones operacionales deben ser considerados como de carácter disposicional. Los estudios lógicos recientes sugieren claramente que pueden definirse las disposiciones con referencia a características manifiestas, como las presentadas por el vocabulario observacional, con la sola ayuda de alguna "modalidad nomológica", tal como el concepto de verdad nomológica, esto es, de verdad en virtud de leyes generales de la naturaleza.[12] Pero presumiblemente un concepto de este tipo es inadmisible según las normas operacionalistas, puesto que no es una característica directamente observable ni definible en términos de tales características.

Otra dificultad se presenta cuando tratamos de dar definiciones completas en términos de observables de expresiones cuantitativas tales como "longitud en centímetros", "duración en segundos", "temperatura en grados Celsius", etc. En la teoría científica se admite que cada una de estas magnitudes asuma cualquier valor real dentro de cierto intervalo; por lo tanto, se plantea el problema de si cada uno de los infinitos valores permisibles, por ejemplo, de la longitud, puede recibir una especificación operacional de significado. Se puede demostrar que es imposible caracterizar cada uno de los valores numéricos permisibles mediante alguna combinación extensional de características observables, ya que la existencia de un umbral de discriminación en todos los ámbitos de observación sólo admite un número finito de combinaciones no equivalentes de este tipo.[13]

[12] Para dar un ejemplo breve, parece razonable *prima facie* definir "x es soluble en agua" por "si se coloca x en agua, entonces x se disuelve". Pero si se interpreta aquí la frase "si..., entonces..." como el condicional extensional o "material", los objetos calificados de solubles por la definición incluyen, entre otros, todas aquellas cosas que nunca se colocan en agua, independientemente de que sean o no realmente solubles en agua. Sólo puede evitarse esta consecuencia, que es un aspecto de las "paradojas de la implicación material", si se concibe nuestro definiens condicional de una manera más restrictiva. Se presenta la idea de interpretar "x es soluble en agua" como una forma abreviada de "en virtud de algunas leyes generales de la naturaleza, x se disuelve si se lo coloca en agua", o más brevemente, "es nomológicamente verdadero que si se coloca x en agua, entonces x se disuelve". La frase "si..., entonces...", puede ahora ser entendida nuevamente en el sentido extensional. Sin embargo, que este análisis sea o no aceptable depende, por supuesto, de si puede considerarse la verdad nomológica como un concepto suficientemente claro. Se encontrará un examen más completo de este complejo problema particularmente en R. Carnap, "Testability and Meaning", *Philosophy of Science*, 3 (1936) y 4 (1937), y N. Goodman, "The Problem of Counterfactual Conditionals", *Journal of Philosophy*, 44 (1947).

[13] En otras palabras, no es posible suministrar para todo valor teóricamente permisible, r, de la longitud $l(x)$ de una barra x una definición de la forma:

$$[l(x) = r] =_{df} C(P_1x, P_2x, \ldots, P_nx)$$

donde P_1, P_2, ..., P_n son características observables, y el definiens es una expresión formada a partir de "P_1x", "P_2x", ..., "P_nx" por medio de los conectivos "y", "o" y "no" solamente.

Cabe señalar, sin embargo, que si las constantes lógicas admitidas en el definiens incluyen, además de conectivos extensionales, cuantificadores y el signo de identidad,

Dificultades como ésta sugieren la idea de si no sería posible concebir métodos más generales y más flexibles que la definición para la introducción de términos científicos sobre la base del vocabulario observacional. Carnap ha elaborado un método semejante que utiliza las llamadas oraciones de reducción. Estas constituyen una versión bastante generalizada de las oraciones definitorias y son particularmente apropiadas para la reformulación precisa del propósito de las definiciones operacionales. Como señalamos antes, una definición operacional del tipo más simple estipula que el concepto que introduce, digamos *C*, debe aplicarse a aquellos casos, y sólo a ellos, que, en condiciones de test específicas, *S*, presentan una cierta respuesta característica, *R*. En la teoría de Carnap, se reemplaza esta estipulación por la oración:

(1) $$Sx \rightarrow (Cx \equiv Rx)$$

o, en palabras: si un caso *x* satisface la condición de test *S*, entonces *x* es un ejemplo de *C* si y sólo si *x* manifiesta la respuesta *R*. La fórmula 1, llamada una oración bilateral de reducción, no es una definición completa (que tendría que ser de la forma "$Cx \equiv \ldots$", donde "Cx" constituye el definiendum); ella no especifica el significado de "Cx" para todos los casos, sino solamente para aquellos que satisfacen la condición *S*. En este sentido, sólo constituye una definición parcial o condicional de *C*.[14] Si "*S*" y "*R*" pertenecen al vocabulario observacional de la ciencia, la fórmula 1 es el esquema del tipo más simple de definición operacional, que alude (casi) exclusivamente a operaciones instrumentales o, mejor dicho, a hallazgos experienciales. Las definiciones operacionales que también utilizan operaciones simbólicas estarían representadas por cadenas de oraciones de reducción que contendrían símbolos lógicos o matemáticos. Algunos de tales símbolos aparecen aun en la fórmula 1, sin em-

---

entonces un vocabulario observacional finito puede permitir la definición explícita de una cantidad infinita numerable de otros términos. Por ejemplo, si se incluyen "*x* contiene espacialmente a *y*" e "*y* es una manzana" en el vocabulario observacional, entonces es posible definir las expresiones "*x* contiene 0 manzanas", "*x* contiene exactamente 1 manzana", "*x* contiene exactamente 2 manzanas", etc., de acuerdo con la definición de Frege-Russell de los números naturales. Sin embargo, aunque se aprueben las definiciones de este tipo —y sin duda se las aprueba, de acuerdo con la intención del análisis operacionalista—, subsisten serios obstáculos para una explicación operacionalista de la totalidad de los números reales admitidos como valores teóricos de la longitud, la masa, etc. Sobre este punto, véase C. G. Hempel, ***Fundamentals of Concept Formation in Empirical Science*** (Chicago, University of Chicago Press, 1952), sección 7.

14 El uso de oraciones de reducción elude una de las dificultades halladas en el intento de dar definiciones explícitas, y por ende complejas, de los términos disposicionales: los signos del condicional y del bicondicional que aparecen en la fórmula 1 pueden ser interpretados extensionalmente sin dar origen a consecuencias indeseables del tipo de las caracterizadas en la nota 11. Para mayores detalles, véase R. Carnap, "Testability and Meaning", secciones 5-10; y también C. G. Hempel, ***Fundamentals of Concept Formation in Empirical Science***, secciones 6 y 8. Dicho sea de paso, en el procedimiento de Carnap no se evita totalmente el uso de conceptos nomológicos; se exige de las oraciones de reducción permitidas para la introducción de nuevos términos, que satisfagan ciertas condiciones de validez lógica o nomológica. Véase R. Carnap, "Testability and Meaning", págs. 442-443.

bargo; y evidentemente no puede haber definición operacional alguna que no utilice en absoluto conceptos lógicos.

4. *Sistemas interpretativos*

Una vez advertida la idea de una especificación parcial de significado, parece innecesariamente restrictivo limitar las oraciones que efectúan tal interpretación parcial a oraciones de reducción en el sentido de Carnap. Podría expresarse una especificación parcial de los significados de un conjunto de términos no observacionales de una manera más general, mediante una o más oraciones que vinculen esos términos con el vocabulario observacional, pero que no tengan la forma de oraciones de reducción. Y parece conveniente, con el mismo fin, favorecer incluso las estipulaciones expresadas por oraciones que sólo contienen términos no observacionales; por ejemplo, la estipulación de que dos términos teóricos deben ser considerados como excluyentes puede ser contemplada como una limitación y, en tal sentido, como una especificación parcial de sus significados.

En general, pues, podría introducirse un conjunto de uno o más términos teóricos, t_1, t_2, ..., t_n, mediante algún conjunto M de oraciones tal que: (I) M no contiene ningún término extralógico fuera de t_1, t_2, ..., t_n, y contiene términos observacionales; (II) M es lógicamente consistente; y (III) M no es equivalente a una verdad de la lógica formal. Las últimas dos condiciones sirven simplemente para excluir casos extremos triviales. Un conjunto M de este tipo será llamado brevemente un *sistema interpretativo,* y sus elementos *oraciones interpretativas.*

Las definiciones explícitas y las oraciones de reducción son tipos especiales de oraciones interpretativas, y lo mismo los postulados del significado propuestos recientemente por Kemeny y Carnap.[15]

Las oraciones interpretativas utilizadas en una teoría determinada pueden ser consideradas simplemente como postulados de esta teoría,[16] y todos los términos observacionales, así como los introducidos por el sistema interpretativo, como primitivos. Concebida de este modo, la especificación de los significados de términos no observacionales de la ciencia se asemeja a lo que se ha llamado a veces la definición implícita de los términos primitivos de una teoría axiomatizada por sus postulados. En este último procedimiento, los términos primitivos son todos no interpretados, y los postulados imponen restricciones a toda interpretación de los términos primitivos que puedan convertir los postulados en oraciones verdaderas. Tales restricciones pueden ser consideradas como especificaciones parciales del significado. Sin embargo, el uso de sistemas interpretativos como los que aquí consideramos tiene esta peculiaridad distintiva: los términos primitivos incluyen un conjunto de términos —los ob-

[15] J. G. Kemeny, "Extension of the Methods of Inductive Logic", *Philosophical Studies,* 3 (1952); R. Carnap, "Meaning Postulates", *ibíd.,* 3 (1952).

[16] Para el caso de las oraciones de reducción de Carnap, la interpretación postulacional me fue sugerida por N. Goodman y A. Church.

servacionales— que se comprenden de antemano y, por ende, no necesitan ninguna interpretación, y con referencia a los cuales los postulados efectúan una especificación parcial de significado para los términos primitivos restantes, los no observacionales. Esta especificación parcial consiste nuevamente en limitar las interpretaciones de los términos no observacionales que hagan verdaderos a los postulados.

5. *Implicaciones para la idea de significado experimental y para la distinción en la ciencia de oraciones analíticas y sintéticas*

Si se concibe de esta manera amplia la introducción de términos no observacionales, lo cual parece coincidir con las necesidades de una reconstrucción formal del lenguaje de la ciencia empírica, entonces es inútil pedir la definición operacional o inquirir por el contenido experiencial de un término teórico. Ya no se dispone en general de la definición explícita por medio de observables, y sólo puede atribuirse significado experiencial —u operacional— al conjunto de todos los términos no observacionales que funcionan en una teoría dada.

Además, no queda ninguna manera general satisfactoria de dividir todos los sistemas concebibles de términos teóricos en dos clases: los que son científicamente significativos y los que no lo son; los que tienen contenido experiencial y los que carecen de él. En cambio, la significación experiencial u operacional parece susceptible de gradación. Para comenzar con una posibilidad extrema: el sistema interpretativo *M* que introduce los términos dados puede ser simplemente un conjunto de oraciones con la forma de definiciones explícitas que provean de un equivalente observacional a cada uno de esos términos. En este caso, los términos introducidos por *M* tendrán una significación experiencial máxima, por así decirlo. En otro caso, *M* puede consistir en oraciones de reducción para los términos teóricos; ellas nos permitirán formular, en términos de observables, una condición necesaria y una condición suficiente (distinta) para la aplicación de cada uno de los términos introducidos. También *M* podría contener oraciones con la forma de definiciones u oraciones de reducción solamente para algunos de los términos no observacionales que introduce. Y por último, podría suceder que ninguna de las oraciones de *M* tuviera la forma de una definición o una oración de reducción; no obstante esto, una teoría cuyos términos sean introducidos por un sistema interpretativo de este tipo puede permitir el testeo mediante hallazgos observacionales, y en ese sentido el sistema de sus términos no observacionales puede poseer contenido experiencial.[17]

[17] Ilustra esto el siguiente modelo simple: la teoría *T* consiste en la oración "$(x)\ ((C_1x \cdot C_2x) \rightarrow C_3x)$" y sus consecuencias lógicas; los tres términos "teóricos" que aparecen en ella son introducidos por el conjunto interpretativo *M* formado por oraciones "$O_1x \rightarrow (C_1x \cdot C_2x)$" y "$(C_1x \cdot C_2x) \rightarrow (O_2x \vee O_3x)$", donde "$O_1$", "$O_2$", "$O_3$" pertenecen al vocabulario observacional. Como es fácil ver, *T* permite en virtud de *M* la "predicción" de que si un objeto tiene la propiedad observable O_1, pero carece de la propiedad observable O_2, entonces tendrá la propiedad observable O_3. De este modo, puede someterse *T* a test experimental, aunque *M* no brinda para ninguno de sus términos constituyentes un criterio de aplicación observacional y operacional que sea necesario y suficiente.

Así, parece posible establecer grados de significación experiencial, y todo intento de trazar una dicotomía que sólo admita sistemas conceptuales experiencialmente significativos y sistemas conceptuales experiencialmente sin significado parece demasiado tosco para adecuarse al análisis lógico de los conceptos y teorías científicos.

Los sistemas interpretativos ofrecen un método más general de introducir términos teóricos que el método de los postulados de significado que elaboraron Carnap y Kemeny. En efecto, aunque los postulados de significado son concebidos como analíticos y, por consiguiente, como implicando sólo consecuencias analíticas, un sistema interpretativo puede implicar ciertas oraciones que contengan términos observacionales pero no términos teóricos y que no sean leyes formales de la lógica ni analíticas en el sentido habitual. Consideremos, por ejemplo, las dos siguientes oraciones interpretativas, que consituyen lo que Carnap llama un par reductivo, y que interpreta "*C*" por medio de predicados observacionales, "R_1", "S_1", "R_2", "S_2":

(2.1) $$S_1x \rightarrow (R_1x \rightarrow Cx)$$

(2.2) $$S_2x \rightarrow (R_2x \rightarrow \sim Cx).$$

Puesto que en ningún caso es posible satisfacer conjuntamente las condiciones suficientes para *C* y para $\sim C$ (*no-C*), las dos oraciones implican la consecuencia [18] de que, para todo caso *x*,

(3) $$\sim (S_1x \cdot R_1x \cdot S_2x \cdot R_2x)$$

Esto es, ningún caso *x* presenta los atributos S_1, R_1, S_2, R_2, juntos. Ahora bien, una aserción de este tipo no es una verdad de la lógica formal, ni puede ser considerada en general como verdadera exclusivamente en virtud del significado de sus términos constituyentes. En consecuencia, Carnap considera esta consecuencia de las fórmulas 2.1 y 2.2 como empírica y como expresión del contenido fáctico del par reductivo a partir del cual fue deducida. Los casos de este tipo no se limitan en modo alguno a las oraciones de reducción, y vemos que en el uso de sistemas interpretativos *la especificación de significado* y *la enunciación de un hecho empírico* —dos funciones del lenguaje que a menudo se consideran completamente distintas— se vinculan tan íntimamente que plantean serias dudas acerca de la conveniencia, o aun la posibilidad, de conservar esta distinción en una reconstrucción lógica de la ciencia. Esta consideración sugiere que abandonemos la distinción, mantenida hasta ahora con fines expositivos, entre las oraciones interpretativas —incluidas en *M*— y el resto de las ora-

[18] Carnap la llama la oración representativa del par de fórmulas 2.1. y 2.2. Véase R. Carnap, "Testability and Meaning", págs. 444 y 451. En general, cuando se introduce un término por varias oraciones de reducción que representan criterios operacionales de aplicación diferentes, entonces el acuerdo entre los resultados de los procedimientos correspondientes, acuerdo que debe presuponerse para que las oraciones de reducción sean compatibles entre sí, están expresados por la oración representativa asociada al conjunto dado de oraciones de reducción. La oración representativa refleja, por lo tanto, el riesgo inductivo que, como ha destacado Bridgman, se corre al usar más de un criterio operacional para un término dado.

ciones que constituyen una teoría científica: podemos simplemente concebir los dos conjuntos de oraciones como una única "teoría interpretada".

Los resultados obtenidos en este breve análisis de la concepción operacionalista de los conceptos científicos significantes son muy semejantes a los que pueden obtenerse mediante un estudio similar de la concepción empirista-lógica de los enunciados o hipótesis científicos significantes.[19] En el último caso, el requisito original de la verificabilidad o la refutabilidad completa por datos experienciales debe ceder ante la exigencia más liberal de la confirmabilidad, esto es, de la verificabilidad parcial. Puede demostrarse que esta exigencia es aplicable propiamente a sistemas teóricos totales, más que a hipótesis individuales, punto ya señalado por Pierre Duhem. Se ve, pues, que la significación experiencial es una cuestión de grado, de modo que la tajante distinción original entre hipótesis (o sistemas de ellas) cognoscitivamente significativas e hipótesis cognoscitivamente sin significado debe ser abandonada. Y hasta parece dudoso que la distinción entre oraciones analíticas y oraciones sintéticas pueda ser mantenida de modo efectivo con referencia al lenguaje de la ciencia empírica.

[19] Véase el ensayo "Criterios empiristas de significación cognoscitiva: problemas y cambios" en este mismo volumen. Sobre la noción de analiticidad, véase W. V. Quine, "Two Dogmas of Empiricism", *Philosophical Review*, 40 (1951).

Tercera Parte

ESTRUCTURA Y FUNCION DE LOS CONCEPTOS Y TEORIAS CIENTIFICOS

CAPÍTULO VI

Fundamentos de la taxonomía

1. *Introducción*

En este artículo [1] trataremos de establecer una base sistemática para el examen de la taxonomía [2] de los desórdenes mentales. Con tal fin, analizaremos los aspectos lógicos y metodológicos básicos de los procedimientos clasificatorios usados en diversas ramas de la ciencia empírica e indicaremos algunas implicaciones que este análisis parece sugerir para los problemas taxonómicos de la psiquiatría.

2. *Clases y conceptos*

Como es bien sabido, una clasificación divide un conjunto o clase dada de objetos en subclases. Los objetos son llamados los *elementos* o *miembros* del conjunto dado; al conjunto mismo lo llamaremos también el *universo del discurso,* en particular cuando se supone que contiene como elementos a todos los objetos de los que se ocupa una investigación determinada.

Los objetos de una clasificación pueden ser cosas concretas, tales como estrellas, cristales, organismos, libros, etcétera; o pueden ser entidades abstractas, tales como números, sistemas de parentesco, ideologías políticas, religiones o doctrinas filosóficas.

Puede concebirse cada una de las subclases establecidas en una cla-

[1] Lo que sigue es la esencia de un trabajo leído ante la Conferencia de Trabajo sobre Estudios de Campo en Desórdenes Mentales, realizada en Nueva York en febrero de 1959, bajo los auspicios de la Asociación Norteamericana de Psicopatología. El texto actual contiene algunos cambios que introduje en la versión original como resultado de la discusión de mi memoria. Las memorias leídas en la Conferencia, a algunas de las cuales aludo mediante los nombres de sus autores, fueron publicadas en Zubin (1961), que también contiene una transcripción de las discusiones.

[2] El término "taxonomía" a menudo es sinónimo de "clasificación", pero aquí usaré las voces "taxonomía" y "taxonómico" fundamentalmente para referirme a la *teoría* de los procedimientos y sistemas clasificatorios. Los dos conceptos distinguidos de este modo han sido caracterizados de manera más completa en el prefacio del estudio de Gregg (1954), donde se contrapone la "taxonomía propiamente dicha" a la "taxonomía metodológica".

sificación dada como definida por la especificación de condiciones necesarias y suficientes para ser miembro de ella, es decir, mediante la enunciación de ciertas características que poseen todos los miembros de esta clase y sólo ellos. Cada subclase queda, pues, definida por medio de (o más precisamente, como una extensión de) un determinado *concepto,* que representa al conjunto de características esenciales para ser miembro de esta subclase. Por ejemplo, en la división de los enteros positivos en números primos y números compuestos, la condición para ser miembro de la primera de estas subclases es que el número en cuestión sea mayor que 1 y sólo sea un múltiplo entero de 1 y de sí mismo. Estas características determinan el concepto de número primo, y la clase correspondiente es la extensión de este concepto.

Análogamente, cada uno de los grupos jerárquicamente ordenados (clases, órdenes, familias, géneros, especies, razas, etcétera) de una clasificación de los mamíferos puede ser considerado como la extensión de un concepto correspondiente, por ejemplo, del concepto de marsupial, murciélago, primate, etcétera.

De igual modo, las subclases establecidas por un sistema taxonómico particular de desórdenes mentales están determinadas por los diferentes tipos de enfermedades mentales que se distingan conceptualmente en el sistema. Por ejemplo, en el sistema del *Manual diagnóstico y estadístico* de la Asociación Psiquiátrica Americana, la especificación del concepto de reacción psicótica depresiva sirve para determinar la clase de los individuos a quienes se aplica el concepto, es decir, que sufren de este tipo de reacción. Como aclara este ejemplo, los objetos de clasificación en la taxonomía psiquiátrica no son los diversos tipos de desórdenes mentales, sino los casos individuales, que son ubicados en diversas clases según el tipo de desorden mental que ejemplifiquen. Esta interpretación concuerda perfectamente con la concepción del *diagnóstico* como la asignación de casos individuales a clases particulares en un sistema taxonómico de enfermedades; y es definidamente requerido por el uso que se hace de las clasificaciones psiquiátricas en las estadísticas médicas, que se ocupan de la distribución de los casos individuales en las diversas clases establecidas en un sistema clasificatorio, como la *Clasificación estadística internacional de las enfermedades* y el *Manual diagnóstico y estadístico.*

Un caso individual del tipo al que aquí aludimos puede ser entendido mejor como un ser humano particular en un momento determinado, o durante un lapso determinado de la historia de su vida; esta interpretación admite la posibilidad de que una persona pueda pertenecer a una clase que representa a una cierta enfermedad en un tiempo determinado, pero no en todo tiempo, de su vida. (En cambio, es mejor considerar los elementos clasificados por un sistema taxonómico de la biología como organismos individuales durante el lapso total de su vida.)

Otras maneras alternativas de dividir un universo dado del discurso en subclases, corresponden al uso de conjuntos alternativos de conceptos para destacar semejanzas y diferencias entre los objetos en estudio. Así, las diferentes tipologías del cuerpo y del temperamento que se han elaborado desde la antigüedad hasta el presente emplean diferentes conjuntos

de conceptos para clasificar o tipificar a una persona determinada. Por ejemplo, uno de los sistemas para clasificar individuos según sus temperamentos se basa en los conceptos de extraversión e introversión; otro, en los de cerebrotonía, viscerotonía y somatotonía; otro, en los conceptos de temperamentos cicloide y esquizoide, etcétera. Los esquemas clasificatorios o tipológicos resultantes difieren según los conceptos adoptados.

Por ende, la especificación de un sistema clasificatorio exige un conjunto correspondiente de conceptos clasificatorios: cada clase establecida en el sistema es la *extensión* de uno de esos conceptos; es decir, consiste exactamente en esos objetos del universo del discurso que poseen las características específicas representadas por el concepto. Por consiguiente, puede considerarse la elaboración de un adecuado sistema de clasificación en un dominio determinado de la investigación como un tipo especial de formación de conceptos científicos. Parece razonable, pues, en un estudio metodológico de la taxonomía, examinar primero las funciones básicas de los conceptos científicos en general, y luego considerar qué requisitos imponen tales funciones a los conceptos clasificatorios.

En nuestro examen, distinguiremos entre los *conceptos* y los *términos* que los representan, como se hace por lo común en la lógica contemporánea. Por ejemplo, el término "soluble en alcohol", que es una expresión lingüística, representa al concepto de solubilidad en alcohol, que es una propiedad de ciertas sustancias. Nos referiremos colectivamente a los términos usados por la ciencia empírica en general o por alguna de sus ramas como a su *vocabulario.*

3. *La descripción y la sistematización teórica como dos funciones básicas de los conceptos científicos*

Hablando en términos amplios, el vocabulario de la ciencia tiene dos funciones básicas: primero, permitir una *descripción* adecuada de las cosas y sucesos que son objeto de la investigación científica; segundo, permitir el establecimiento de leyes o teorías generales mediante las cuales sea posible *explicar y predecir* —y por ende, *comprender científicamente*— sucesos particulares. En efecto, comprender científicamente un fenómeno es demostrar que se produce de acuerdo con leyes generales o principios teóricos.

Con alguna simplificación, a menudo puede decirse que el desarrollo de una disciplina científica parte de una etapa inicial de la "historia natural",[3] etapa en la que se busca primeramente describir los fenómenos en estudio y establecer generalizaciones empíricas simples acerca de ellos, para pasar a etapas cada vez más "teóricas", en las cuales se otorga creciente importancia al logro de explicaciones teóricas amplias del tema empírico en investigación. El vocabulario requerido en las primeras etapas de este desarrollo será en su mayor parte observacional: se lo elegirá de modo que permita la descripción de esos aspectos del objeto de estudio

[3] Este sugerente término está tomado de Northrop (1947), en particular de los capítulo 3 y 4, donde se establece una distinción entre "la etapa de historia natural de la investigación" y la "etapa de la teoría formulada deductivamente".

que sean discernibles de manera bastante directa mediante la observación. El paso a la sistematización teórica se señala por la introducción de nuevos términos "teóricos", que se refieren a las diversas entidades postuladas teóricamente, a sus características y a los procesos en que intervienen; todos ellos están más o menos lejos del nivel de las cosas y los sucesos directamente observables. Por ejemplo, los campos eléctricos y magnéticos de la física y la propagación de ondas en ellos, las valencias químicas, las estructuras moleculares y atómicas, las partículas elementales de la física, los estados cuánticos, etcétera, son casos típicos de los tipos de cosas y procesos a los que se refiere el vocabulario teórico de la física y de la química.

En la medicina, el paso de una etapa predominantemente descriptiva a otras cada vez más teóricas se refleja, por ejemplo, en la transición de un punto de vista principalmente sintomatológico a otros cada vez más etiológicos. No debe concebirse la etiología como el estudio de las "causas" de las enfermedades en un sentido estrecho de este término. En las ciencias físicas, la búsqueda de causas en tal sentido ha sido reemplazada por la búsqueda de leyes y teorías explicativas; y la ideología se ha desplazado en la misma dirección. En verdad, los diversos enfoques teóricos de las enfermedades han dado origen a una variedad de conceptos teóricos. Por ejemplo, el *Manual diagnóstico y estadístico* caracteriza el concepto de reacción de conversión del siguiente modo:

> En lugar de ser experimentado conscientemente..., el impulso que causa la ansiedad es "convertido" en síntomas funcionales en los órganos o parte del cuerpo, habitualmente en aquellos que se encuentran sobre todo bajo control voluntario. Los síntomas sirven para disminuir la ansiedad consciente (sentida) y por lo común son típicos del conflicto mental subyacente. Tales reacciones casi siempre satisfacen las necesidades inmediatas del paciente y están, por ende, asociadas a una "ganancia secundaria" más o menos obvia. (págs. 32-33).

Evidentemente, varios de los términos usados en este pasaje no se refieren a fenómenos directamente observables, como la conducta manifiesta, ni a respuestas que puedan ser provocadas por estímulos adecuados, sino más bien a factores psicodinámicos postulados teóricamente. Esos términos tienen un significado y una función claros sólo en el contexto de su teoría correspondiente, así como los términos "campo gravitacional", "potencial gravitacional", etcétera, sólo tienen un significado y una función definidos en el contexto de su teoría correspondiente de la gravitación.

Revisemos ahora algunos de los requisitos que los dos objetivos fundamentales de la descripción y la sistematización teórica imponen a los conceptos científicos, y en particular a los conceptos usados con fines clasificatorios.

4. *Contenido empírico de los términos científicos: definición operacional*

La ciencia aspira al conocimiento *objetivo*, en el sentido de ser intersubjetivamente certificable, independientemente de la opinión o preferencia individual, sobre la base de datos asequibles mediante experimentos u observaciones apropiados. Esto exige que los términos usados para formular enunciados científicos tengan significados claramente especificados y sean entendidos en el mismo sentido por todos aquellos que los usan. Una de las principales objeciones contra diversos tipos de teorías psicodinámicas contemporáneas, por ejemplo, es que sus conceptos fundamentales carecen de criterios claros y uniformes de aplicación, y que, en consecuencia, no hay maneras definidas e inequívocas de someter las teorías a test aplicándolas a casos concretos.

Un método que se ha recomendado mucho para evitar este tipo de deficiencias es el uso de las llamadas *definiciones operacionales* de los términos científicos. La idea fue expuesta muy explícitamente por vez primera por el físico P. W. Bridgman en su libro *La lógica de la física moderna.* Se piensa que una definición operacional de un término dado brinda criterios objetivos por medio de los cuales cualquier investigador científico puede decidir, en cualquier caso particular, si el término tiene o no aplicación. Con tal fin, la definición operacional especifica una "operación" de testeo, *T*, que puede ser llevada a cabo en cualquier caso al cual el término dado pueda aplicarse presumiblemente, y un cierto resultado *R* de la operación de testeo cuya aparición es considerada como el criterio para la aplicabilidad del término al caso dado. Esquemáticamente, una definición operacional de un término científico *C* es una estipulación según la cual *C* debe aplicarse a todos aquellos casos, y sólo a ellos, para los que la realización de la operación de testeo *T* da el resultado especificado *R.* Como ejemplo: una definición operacional simple del término *más duro que* tal como se lo usa en mineralogía podría especificar que un trozo de mineral *x* será dicho más duro que otro trozo de mineral *y* si la operación de deslizar una punta de *x* bajo presión sobre una superficie lisa de *y* da como resultado una rayadura en *y*, mientras que *y* no raya de igual modo a *x*. Análogamente, una definición operacional de longitud debe especificar reglas para la medición de longitud en términos de operaciones públicamente realizables, tales como el uso apropiado de varas de medir. De igual modo, podría definirse operacionalmente la oligofrenia fenilpirúvica con referencia a la "operación" de testear químicamente la orina de la persona implicada para determinar la presencia de ácido fenilpirúvico; el "resultado" que indica la presencia de la condición (y por ende la aplicabilidad del término correspondiente) es simplemente un resultado positivo del test. La mayor parte de los procedimientos de diagnóstico usados en la medicina se basan en criterios operacionales de aplicación para las correspondientes categorías de diagnóstico. Pero hay excepciones. Por ejemplo, se ha sugerido que la aparición de un "sentimiento precoz" característico en el investigador puede considerarse como un indicio de demencia precoz en el paciente que está examinando: pero esta idea no cumple con los requisitos del operaciona-

lismo, porque la aparición del resultado especificado, el sentimiento precoz con respecto a un paciente determinado, *no* es independiente del examinador.

Bridgman arguye, en efecto, que si los significados de los términos usados en una disciplina científica son especificados operacionalmente, entonces las aserciones de esta disciplina pueden recibir un testeo objetivo. En cambio, si un problema o hipótesis propuestos están expresados en términos tales que algunos de ellos no tienen tal vínculo con la tierra firme de los datos asequibles operacionalmente, el operacionalismo los rechaza y los considera como carentes científicamente de significado, porque no pueden recibir apoyo de ningún test empírico, de modo que la formulación propuesta, a su vez, no puede tener vinculación alguna con asuntos empíricos, y por ende carece de contenido empírico.[4] La insistencia operacionalista en que los términos científicos significativos tengan criterios definidos y públicos de aplicación presenta, pues, una estrecha afinidad con la insistencia empirista en que las hipótesis y teorías científicas significativas puedan recibir, en principio, un test intersubjetivo mediante datos observacionales.

Las tesis metodológicas del operacionalismo y el empirismo han hallado un interés especialmente agudo y en gran medida favorable en la psicología y la sociología. En ellas, a menudo se logra una especificación operacional del significado formulando procedimientos de testeo definidos que deben regir la aplicación de términos tales como "*CY*" y de términos referentes a diversas aptitudes y actitudes.

La preocupación de muchos psicólogos y científicos sociales por la *confiabilidad* de sus términos refleja la importancia atribuida a la objetividad del uso: habitualmente se interpreta la confiabilidad de un concepto (o del término correspondiente) como indicio de dos cosas: la coherencia manifestada en su uso por un observador, y el acuerdo en el uso que hacen de él diferentes observadores. A menudo se expresa el primer aspecto en términos de la correlación entre los juicios hechos por el mismo observador cuando se le pide que juzgue el mismo caso en ocasiones diversas; el segundo aspecto se expresa en términos de las correlaciones existentes entre los juicios de varios observadores que juzgan los mismos casos; los "juicios" aquí aludidos se realizan en términos del concepto cuya confiabilidad está en consideración.

La importancia que otorgan los operacionalistas a la determinación de criterios públicos claros y precisos de aplicación de los términos científicos es sin duda correcta y saludable. Pero las formulaciones habituales del operacionalismo exigen ciertas reservas, dos de las cuales serán mencionadas aquí brevemente porque se relacionan con el tema de este artículo.

En primer lugar, los criterios operacionales de aplicación de un término a menudo equivalen a algo menos que una definición completa. Por ejemplo, pueden especificarse criterios de aplicación para el término *temperatura* con referencia a la operación de colocar un termómetro de mercurio en el lugar apropiado y observar su respuesta; o por un uso

[4] Véase, por ejemplo, Bridgman, pág. 28.

similar de un termómetro de alcohol, o de una termocupla, etcétera. Estos instrumentos tienen ámbitos diferentes, aunque se superponen parcialmente, dentro de los cuales pueden ser usados, y ninguno de ellos abarca toda la gama de las temperaturas teóricamente posibles. Cada uno de ellos suministra, pues, una *definición parcial* o mejor dicho un *criterio parcial de aplicación* del término considerado (o del concepto correspondiente). Tales criterios parciales de aplicación de los términos que aparecen en una hipótesis o teoría determinada a menudo bastan para permitir un test empírico. En verdad, hay razones para dudar de la posibilidad de establecer definiciones operacionales *completas* para todos los términos teóricos de la ciencia, por lo cual el programa operacionalista debe hacerse más liberal, de modo que exija solamente la especificación de criterios parciales de aplicación.[5]

En segundo lugar, para que la insistencia en una especificación *operacional* de significado para términos científicos no sea indebidamente restrictiva, debe tomarse la idea de operación en un sentido muy liberal que no exige la manipulación de los objetos implicados: la mera observación de un objeto, por ejemplo, debe ser considerada como una operación. En efecto, los criterios de aplicación de un término pueden ser especificados con referencia a ciertas características que es posible discernir sin ningún procedimiento de testeo más complicado que la observación directa. Consideremos, por ejemplo, la lista de características que da Sheldon para el endomorfismo dominante. Esta lista incluye características observables directamente tales como la redondez y suavidad del cuerpo, la concentración central de la masa, los hombres altos y cuadrados y de suaves contornos, el cuello corto y los miembros cortos y ahusados.[6] Esta es una manera satisfactoria de determinar el concepto de endomorfismo predominante y, por consiguiente, la clase de los individuos predominantemente endomórficos, siempre que los términos usados para especificar las carateristicas distintivas de las endomorfías tengan un significado razonablemente preciso y sean utilizados por todos los investigadores que intervienen con un alto grado de uniformidad intersubjetiva; es decir, siempre que, para todo sujeto dado, haya un alto grado de acuerdo entre los diferentes observadores acerca de si dicho sujeto tiene o no suaves contornos corporales, cuello corto, miembros ahusados, etcétera. Y en verdad, la insistencia de Bridgman en los tests operacionales y sus resultados tiende, sin duda, a asegurar que los criterios de la aplicación de conceptos científicos se expresen en términos con un grado muy elevado de uniformidad en su uso.

Pero no sería razonable exigir que *todos* los términos usados en una disciplina científica determinada reciban una especificación operacional de significado, pues en tal caso, el proceso de especificar los significados de los términos definitorios, etcétera, conduciría a un regreso infinito. En todo contexto definicional (independientemente del problema del ope-

[5] Se hallará un examen más detallado de estos problemas en Hempel (1958).

[6] Ver Sheldon, Stevens y Tucker (1940), pág. 37. Para la somatotipia detallada se necesita una serie de diámetros de la superficie corporal y, por ende, la "operación" de aplicar aparatos de medición adecuados; véase *loc. cit.*, cap. 3.

racionalismo), algunos términos deben ser comprendidos de antemano; y la objetividad de la ciencia exige que los términos que sirven de base a la introducción de otros términos científicos se cuenten entre los usados con un alto grado de uniformidad por diferentes investigadores del mismo campo.

Justamente por esta razón, habitualmente se formulan los criterios operacionales de aplicación de términos psicológicos con referencia a aspectos públicamente observables de la conducta que manifiesta un sujeto en respuesta a estímulos específicos públicamente observables, y ésta parece en verdad la manera más satisfactoria de cumplir con las exigencias de la objetividad científica. La referencia a "operaciones" de un carácter altamente introspectivo y subjetivo no satisface los requisitos de la formación de conceptos científicos. Por ejemplo, la reformulación operacional de conceptos psicoanalíticos propuesta por Ellis,[7] que se basa en "operaciones" tales como pensar, recordar, actuar de manera emotiva y percibir (en un sentido enormemente amplio), no brinda ningún criterio claro de aplicación de los términos del psicoanálisis y ninguna manera objetiva de testear las hipótesis psicoanalíticas.

Apliquemos las consideraciones anteriores a la taxonomía de los desórdenes mentales. Si un esquema clasificatorio es usado con un alto grado de uniformidad por diferentes investigadores, los conceptos que determinan las diversas subclases deberán poseer criterios claros de aplicación que puedan ser enunciados en términos de características públicamente discernibles. Opino que la importancia de los criterios objetivos de clasificación, o de criterios objetivos de diagnóstico, ha quedado notablemente ejemplificada por las observaciones hechas en algunos de los otros artículos preparados para esta conferencia. Por ejemplo, el profesor Stengel[8] menciona en su memoria que, entre los casos recibidos en hospitales psiquiátricos de Inglaterra y Gales durante 1949, una fracción improbablemente pequeña fue asignada a las categorías 315-317 (psiconeurosis con síntomas somáticos) de la *Clasificación estadística internacional de las enfermedades*; y se plantea el problema de saber si la falta de criterios claramente especificados de aplicación no puede explicar en parte esta aparente anomalía. Otro caso que viene a propósito es la observación del profesor Greenberg de que no pocas veces los técnicos, ayudantes y hasta los coinvestigadores empeñados en un proyecto de investigación común difieren en sus interpretaciones de los significados de términos, condiciones patológicas y procedimientos, cuando no se les especifica por escrito. En un espíritu similar, el profesor Strömgren señala que muchas de las controversias entre los investigadores de la demografía psiquiátrica pueden atribuirse fácilmente a discrepancias en las definiciones. Pero si bien la formulación de criterios más confiables de aplicación es ciertamente muy deseable, no estoy seguro de que sea siempre una tarea fácil. El profesor Strömgren da algunos ejemplos de esto en su artículo. Sería, pues, poco razonable y frustrante insistir desde el principio en que se adopten las más elevadas normas de precisión. Pero es impor-

[7] Véase Ellis (1956).
[8] Esta contribución y otras que pronto citaremos están en Zubin (1961).

tante tender al logro de criterios cada vez más confiables de aplicación para las diversas categorías establecidas en una clasificación de los desórdenes mentales.

En pro de este objetivo, quizá valga la pena considerar si en la especificación de conceptos psiquiátricos se usan criterios con resonancias valorativas, y si es así, en qué medida. Tomemos, por ejemplo, la caracterización de la categoría "personalidad inadecuada", tal como la da el *Manual diagnóstico y estadístico* (pág. 35): "tales individuos se caracterizan por una respuesta inadecuada a las exigencias intelectuales, emocionales, sociales y físicas. Al examinarlos, no son muy deficientes física ni mentalmente, pero manifiestan inadaptabilidad, ineptitud, pobre capacidad de juicio, carencia de vigor físico y emocional, e incompatibilidad social." Nociones tales como las de inadecuación de respuesta, inadaptabilidad, ineptitud y pobre capacidad de juicio tienen evidentemente aspectos valorativos y es de esperar que su uso en casos concretos reciba la influencia de las idiosincrasias de los investigadores; esto reduce la confiabilidad de esos conceptos y de aquellos para los cuales sirven como criterios parciales de aplicación.

Una manera interesante de aumentar la uniformidad en el uso intersubjetivo de ciertos términos clasificatorios ha sido señalada por Lazarsfeld y Barton: algunos tipos de juicios clasificatorios se hacen más confiables cuando los "indicadores", los criterios que sirven para asignar casos individuales a clases específicas, se descomponen en diversos componentes. Por ejemplo, cuando diferentes clasificadores juzgan la adaptación de los niños, la confiabilidad aumentará simplemente especificando ciertos aspectos a los cuales los clasificadores deben prestar atención, tales como la apariencia (la cual, a su vez, puede ser adicionalmente caracterizada mediante indicadores tales como un cabello y una vestimenta excesivamente desordenados, las uñas comidas, una expresión facial rígida, etcétera), la respuesta a las entrevistas y la actitud hacia otros y hacia sí mismo. Los autores agregan, significativamente, que a pesar del aumento de objetividad que se logra de este modo, "se requiere, con todo, cierto conjunto de enseñanzas y experiencias comunes, como las que manifiestan los psicólogos experimentados de niños, para hacer bien una vaga labor de procedimientos".[9]

Otro factor que puede influir sobre la confiabilidad de los criterios clasificatorios lo ejemplifican el Test de Rorschach, el Test de apercepción temática y procedimientos similares, todos los cuales pueden ser considerados como fuentes de criterios operacionales con propósito de diagnóstico. Estos tests difieren, por ejemplo, de los tests de inteligencia o de aptitudes del tipo común que exigen una parte considerable de interpretación, y que no tienen un procedimiento simple —realizable en principio por una máquina, por así decir— para observar las respuestas del sujeto y combinarlas en un diagnóstico inequívoco que asigne dicho sujeto a alguna clase particular.

Observaciones similares se aplican a la tipología de temperamentos de Sheldon. Para la asignación diagnóstica de un sujeto a uno de los di-

[9] Véase Lazarsfeld y Barton (1951), sobre todo págs. 166-167.

versos tipos establecidos en el sistema, el examinador debe analizar al sujeto con respecto a una lista específica de características; y si bien es probable que haya estrecho acuerdo entre los análisis realizados por examinadores diferentes, Sheldon y Stevens [10] agregan en este comentario sobre el procedimiento:

> El último uso (diagnóstico) de las características, considerando a éstas individualmente, es quizá tan objetivo y sistemático como el diagnóstico médico. Es decir, admitimos que se halla presente un elemento subjetivo, que no se ha construido ninguna máquina que pueda realizar el diagnóstico del temperamento.

Sin embargo, la objetividad, o intersubjetividad en discusión, es cuestión de grados, y debe recordarse que también los resultados de "operaciones" tales como observar un objeto con el microscopio o el telescopio, o un pulmón con el fluoroscopio o indirectamente a través de una fotografía de rayos X, presentan variaciones intersubjetivas aun entre observadores expertos.[11] Lo importante, creo, es ser conscientes de la medida en que los factores subjetivos entran en la aplicación de un conjunto dado de conceptos, y tratar de reducir gradualmente su influencia.

5. *Contenido sistemático y clasificación "natural"*

Pero los criterios claros y objetivos de aplicación no bastan: para ser científicamente útil un concepto debe prestarse a la formulación de leyes generales o principios teóricos que reflejan uniformidades en el objeto de estudio, y de este modo brinden una base para la explicación, la predicción y en general la comprensión científica. A este aspecto de un conjunto de conceptos científicos lo llamaremos su *contenido sistemático,* pues representa la contribución que hacen los conceptos a la sistematización del conocimiento en ese campo por medio de leyes o teorías.

El requisito del contenido sistemático también se aplica, en particular, a los conceptos que determinan clasificaciones científicas. En realidad, puede elucidarse bien la vaga distinción familiar entre clasificaciones "naturales" y "artificiales" como aludiendo a la diferencia entre clasificaciones que son científicamente fructíferas y las que no lo son: en una clasificación del primer tipo, las características de los elementos que sirven como criterio de pertenencia a una clase dada están asociadas, universalmente o con elevada probabilidad, a conjuntos más o menos amplios de otras características. Por ejemplo, los dos conjuntos de características sexuales primarias que determinan la división de los seres humanos en varones y mujeres están asociados cada uno de ellos, mediante leyes generales o conexiones estadísticas, a una gran variedad de rasgos físicos, fisiológicos y psicológicos concomitantes. Es comprensible que una clasificación de este género sea considerada como algo que tiene una existencia

[10] Sheldon y Stevens (1942), pág. 426.

[11] Véase el cap. 1 de Hanson (1958) para un instructivo examen de la visión y la observación científicas como tareas "cargadas de teorías".

objetiva en la naturaleza, que "divide a la naturaleza por sus articulaciones", a diferencia de las clasificaciones "artificiales", en las cuales las características definitorias tienen pocos vínculos explicativos o predictivos con otros rasgos; es el caso, por ejemplo, de la división de los seres humanos en aquellos que pesan menos de cincuenta kilos y todos los demás. (Esto no equivale a negar que la última distinción, así como otras distinciones similarmente "artificiales", puede ser muy útil para ciertos fines prácticos especiales, como, por ejemplo, la clasificación de las impresiones digitales para la identificación de individuos, aunque el contenido sistemático del sistema parezca muy pequeño.)

Análogamente, como señaló W. S. Jevons (antes de que se diera a conocer el sistema periódico), los elementos potasio, sodio, cesio, rubidio y litio, agrupados en la clase de los metales álcali, tienen muchas características comunes: se combinan enérgicamente con el oxígeno, se descomponen en agua a diversas temperaturas y forman óxidos firmemente básicos y altamente solubles en agua; sus carbonatos son solubles en agua, etcétera.[12] Quizás el ejemplo más notable de una clasificación que refleja leyes generales es el sistema periódico de los elementos, sobre el cual basó Mendeleiev un conjunto de predicciones muy específicas, que fueron confirmadas de modo impresionante por las investigaciones posteriores. Como resultado de avances más recientes, se ha dado al sistema, en una forma un tanto modificada, un fundamento teórico más profundo, al demostrarse que refleja, en las clases representadas por las columnas de la tabla periódica, ciertas semejanzas y diferencias en la estructura atómica de los elementos.

Una evolución similar se ha producido en los métodos taxonómicos de la biología. Aun en los primeros sistemas taxonómicos, que se basaban en características más o menos observables en forma directa (principalmente morfológicas), cada clase representa, por supuesto, un conjunto grande de aspectos empíricos asociados; pero como resultado de la teoría de la evolución, se reemplazó la base morfológica de la clasificación por otra más profundamente impregnada por la teoría, a saber, una base filogenética. Las diversas especies, por ejemplo, son "definidas teóricamente, al menos en principio, en términos filogenéticos y genéticos",[13] y las características morfológicas brindan ahora simplemente los criterios observacionales para la distribución de los individuos en especies concebidas en términos filogenéticos.

En la investigación psicológica y psicopatológica, los sistemas tipológicos de Kretschmer[14] y de Sheldon y sus colaboradores, para mencionar dos ejemplos característicos, revelan el intenso interés por conceptos que reflejen uniformidades empíricas y asociaciones estadísticas. En el siste-

[12] Jevons (1877), pág. 675. Véase también el aclarador examen general del cap. 30 de su libro.

[13] Simpson (1945), pág. 13. Véase también la lúcida exposición del mismo tema en el cap. 19, "Los principios de la clasificación", de Simpson, Pittendrigh y Tiffany (1957). En lo concerniente al contenido sistemático de los conceptos clasificatorios de la taxonomía biológica, véanse los ensayos de Huxley y Gilmour en Huxley (1940).

[14] Véase Kretschmer (1925).

ma de Sheldon, las tres "componentes primarias del temperamento" —la viscerotonía, la cerebrotonía y la somatotonía— son caracterizadas por medio de tres grupos correspondientes de características que fueron seleccionadas sobre la base de una intensa aplicación del método del ensayo y del error empírico, de tal manera que las características de cada grupo se correlacionen positivamente entre sí y presenten una correlación relativa con todas o casi todas las características de los otros grupos.[15] Además, una de las principales razones que se esgrimen para atribuir significación científica al sistema se basa en la correlación entre las tres componentes del temperamento, por una parte, y diversas características psicológicas y somáticas, por la otra.

Con respecto a las últimas, se indican ciertas conexiones estadísticas entre las componentes básicas del temperamento y las componentes básicas del organismo —la endomorfía, la ectomorfía y la mesomorfía— establecidas en la teoría de Sheldon de los tipos somáticos.[16] La tipología del carácter y del organismo creada por Kretschmer tiene objetivos similares; y ambos sistemas tratan de mostrar algunas conexiones entre características somáticas y la disposición a ciertos tipos de perturbaciones mentales. Sean cuales fueren los méritos que puedan demostrarse de éstos y otros sistemas similares, los mencionamos aquí como ejemplos de un deliberado esfuerzo por elaborar sistemas clasificatorios (más precisamente, tipologías, en el sentido que examinaremos en la sección siguiente), cuya base conceptual tiene un definido contenido sistemático.

De acuerdo con los requisitos del contenido sistemático, los conceptos usados en un campo determinado de la investigación científica cambiarán con los avances sistemáticos realizados en ese campo: la formación de conceptos marchará a la par de la formulación de leyes y, eventualmente, teorías. Como señalamos antes, puede que las leyes en un comienzo expresen simples conexiones uniformes o estadísticas entre observables; en tal caso, serán formuladas en términos del vocabulario observacional de la disciplina a que pertenecen. Pero un mayor progreso sistemático exigirá la formulación de principios expresados en términos teóricos que se refieran a diversos tipos de entidades inobservables y a sus características. En el curso de tal desarrollo, las clasificaciones definidas con referencia a características manifiestas y observables cederán cada vez más el lugar a sistemas basados en conceptos teóricos. Ejemplifica este proceso el paso de una caracterización y clasificación fenoménico-observacional de los elementos y compuestos químicos y modos teóricos de definirlos y diferenciarlos con referencia a sus estructuras atómicas y moleculares. Para ser inequívocamente aplicables a casos concretos, los conceptos especificados de manera teórica deben poseer, por supuesto, claros criterios empíricos u "operacionales" de aplicación; pero ya no se podrá considerar a éstos como sus características definitorias: el resultado especificado del test operacional sólo constituye un *síntoma* fácilmente observable de la presencia de los aspectos o procesos representados por los conceptos

[15] Véase Sheldon y Stevens (1942), cap. 2.
[16] Véase Sheldon, Stevens y Tucker (1940); en particular el cap. 7; y Sheldon y Stevens (1942), cap. 7.

teóricos. Los "significados" de estos últimos no son cabalmente reflejados por criterios de aplicación (diagnóstico) solamente, sino también —en importante medida— por el sistema teórico al cual pertenecen.

El énfasis dado al contenido sistemático en la formación de conceptos se ha puesto claramente de manifiesto en la elaboración de sistemas clasificatorios de los desórdenes mentales. Ya no se definen ahora los conceptos que determinan las diversas clases de categorías sólo en términos de síntomas, sino más bien en términos de los conceptos fundamentales de *teorías* que pretenden *explicar* la conducta observable, incluso los síntomas, del mismo modo que la teoría molecular y atómica explica las características más directamente observables que hacían las veces de características definitorias en una etapa anterior de la formación de conceptos en la química. Ejemplifican bien esta tendencia varias de las características de desórdenes mentales expuestas en el *Manual diagnóstico y estadístico*, donde se combina una enumeración de ciertos síntomas con una explicación etiológica o de carácter teórico general: las caracterizaciones de las diversas categorías de desórdenes psiconeuróticos (págs. 31-34 del *Manual*) son casos claros.

En un sistema clasificatorio con una base teórica, pues, dos individuos con síntomas similares pueden ser ubicados en clases diferentes, ya que algunos de los tipos de perturbaciones mentales que se diferencian en el nivel etiológico-teórico pueden coincidir parcialmente en los síndromes asociados a ellas, así como dos compuestos químicos diferentes pueden tener en común diversas características directamente observables. De modo análogo, en los sistemas taxonómicos de la biología que tienen una base filogenético-evolutiva, dos especímenes fenoménicamente muy similares pueden ser ubicados en especies muy alejadas una de otra en jerarquía evolutiva, como las especies lobo (*Canis*) y lobo tasmaniano (*Thylacinus*).[17]

Las consideraciones anteriores tienen cierta relación con la cuestión de si el pronóstico de perspectivas y las posibilidades terapéuticas pueden —o quizás hasta deben— incluirse propiamente entre las características definitorias de las enfermedades mentales. Ciertamente, es concebible que se usen conceptos referentes a desórdenes mentales en un contexto teórico que implique determinados pronósticos, y en verdad cabe esperar que así suceda, como resultado de la ulterior investigación. En este caso, podrían definirse los conceptos aludidos dentro del marco de la teoría, mediante propiedades tales que algunas de ellas tengan el carácter de un pronóstico. En cambio, sería contrario a los propósitos prácticos del diagnóstico y la terapia que los criterios operacionales de aplicación de esos conceptos —es decir, los criterios que son la base del diagnóstico médico— exigieran posponer el diagnóstico hasta después de que la enfermedad recorriera todo su ciclo. Para satisfacer estas necesidades prácticas, pues, será menester formular los criterios de aplicación en términos de características más o menos inmediatamente discernibles. Para mencionar un caso similar tomado de la física sería lamentable que la aplicación

[17] Para éste y otros ejemplos, véase el cap. 19 de Simpson, Pittendrigh y Tiffany (1957).

del término *radio* dependiera del criterio según el cual la vida media de ese elemento es de aproximadamente 1800 años; aunque esta vida media es, por cierto, una característica importante del radio.

Debemos observar, sin embargo, que la distinción aquí supuesta entre criterios de aplicación que incluyen pronósticos y criterios que *no* los incluyen es una cuestión de grado. Las definiciones operacionales, por ejemplo, implican pronósticos condicionales concernientes al resultado de ciertas operaciones de test: si *x* es un trozo de mineral más duro que *y*, entonces el test del rayado dará como resultado una rayadura sobre la superficie de *y*; si a través de este alambre pasa una corriente de 1 amperio, la aguja de un amperímetro conectado con él de la manera apropiada señalará la marca correspondiente; etcétera. Análogamente, la prueba de Schick, que suministra un criterio operacional de aplicación del concepto de inmunidad a la difteria, supone un pronóstico de corto alcance concerniente a una reacción cutánea. Y en ciertos casos, puede apelarse a la respuesta a formas particulares de terapia como un criterio basado en un diagnóstico. Pero parece razonable esperar que los progresos que se realicen en la comprensión teórica brinden de manera creciente explicaciones etiológicas y estructurales de las enfermedades físicas y mentales, y que tales explicaciones, a su vez, impliquen criterios basados en diagnósticos y formulados en términos de condiciones, antecedentes o en características físicas o mentales rápidamente discernibles.

Es muy probable, creo, que las clasificaciones de los desórdenes mentales reflejen cada vez más consideraciones teóricas. No me corresponde especular acerca de la dirección que pueden tomar los avances teóricos en este campo ni, en particular, acerca de si las principales teorías serán formuladas en términos biofisiológicos o bioquímicos, o más bien en términos psicodinámicos que carezcan de una interpretación fisiológica o fisioquímica general. Pueden satisfacer los requisitos básicos de las teorías científicas los sistemas teóricos de uno u otro tipo. Para resumir, y haciendo un esbozo esquemático, tales requisitos exigen: 1) una especificación clara de los conceptos básicos usados para representar las entidades teóricas (objetos, estados, procesos, características, etcétera), en función de los cuales la teoría se propone interpretar y explicar los fenómenos empíricos que caen dentro de su dominio de investigación; 2) un conjunto de suposiciones teóricas (leyes básicas, hipótesis fundamentales) formuladas en términos teóricos y que afirman ciertas interrelaciones de las entidades teóricas correspondientes; 3) una interpretación empírica de la teoría, que podría adoptar la forma de criterios operacionales para los términos teóricos o, con mayor generalidad, de un conjunto de leyes de naturaleza estadística o estrictamente universal, que vinculen las características, estados o procesos teóricos con fenómenos observables; 4) la testabilidad en principios de la teoría especificada de este modo; es decir, la teoría y su interpretación deben implicar afirmaciones deductivas o inductivamente definidas acerca de fenómenos observables que deben suceder en condiciones de test especificables, si la teoría es correcta; la producción o no producción de estos fenómenos suministrará, pues, los elementos de juicio confirmatorios o desconfirmatorios de la teoría. Si

una teoría propuesta no tiene en absoluto implicaciones, evidentemente no guarda relación alguna con asuntos empíricos y no puede aspirar a que se la considere como una teoría significativa de la ciencia empírica (ni siquiera como una teoría endeble o falsa, pues estos últimos atributos suponen un conflicto entre la teoría y elementos de juicio experimentales u observacionales atinentes a ella).[18]

Este requisito de la testabilidad con referencia a fenómenos observables descarta, por ejemplo, la concepción neovitalista de los procesos biológicos como determinados, al menos en parte, por fuerzas vitales o entelequias; pues las enunciaciones dadas de esta concepción no tienen implicaciones experimentalmente testeables.

6. *De los conceptos clasificatorios a los comparativos y cuantitativos*

Si bien no es posible predecir los cambios de contenido que sufrirán los conceptos y las teorías acerca de los desórdenes mentales como consecuencia de la investigación ulterior, creo que se pueden prever ciertos cambios en su carácter lógico. En esta sección final, trataré de indicar brevemente la naturaleza de esos cambios.

Hablando en términos estrictos, la clasificación es un asunto de *sí-o-no*, de *o bien*..., *o bien*: se determina una clase A por algún concepto que representa sus características definitorias, y un objeto dado cae en esta clase o fuera de ella según que tenga las características definitorias o carezca de ellas.

Pero en la investigación científica a menudo se encuentra que los objetos se resisten a un rígido encasillamiento de este tipo. Más precisamente: aquellas características del objeto de estudio que, en el contexto dado de investigación se presentan como una fructífera base de clasificación, con frecuencia no pueden ser tratadas como propiedades que un objeto dado posee *o* no posee; más bien son características susceptibles de graduación y que, por ende, un objeto dado puede presentar *más o* menos marcadamente. Como resultado de ello, algunos de los objetos en estudio presentarán al investigador casos límites, que no entran inequívocamente en uno u otro de varios compartimientos limitados de manera clara, sino que exhiben en cierta medida las características de clases diferentes. Por ejemplo, el profesor Strömgren se refiere en su memoria a las dificultades de hallar una línea limítrofe natural que separe todo el grupo de neurosis y psicopatías de lo que no pertenece a él, y señala que las transiciones son graduales en todas las direcciones. Las tipologías del cuerpo y del temperamento brindan otro buen ejemplo, en el cual el carácter gradual de las transiciones ha recibido recientemente cierta atención metodológica especial. Quienes han propuesto sistemas tipológicos a menudo han destacado que raramente se encuentran en la experiencia —si es que los encuentra alguna vez— casos "puros" de los tipos básicos que distinguen, y que los individuos concretos por lo común presentan mez-

[18] Se encontrará una exposición más completa de estos requisitos fundamentales y un análisis crítico de algunas de sus consecuencias en Hempel (1952), (1958).

clas de diversos tipos. A veces los tipos básicos adquieren el carácter de puntos ideales de referencia que marcan, por así decir, los extremos de una escala a lo largo de la cual se ordenan los casos concretos. Así, Kretschmer [19] declara:

> Nunca, ni siquiera en los casos más definidos, damos con un ejemplo puro, en el más estricto sentido de la palabra, sino siempre con casos individuales peculiares de un tipo, esto es, el tipo mismo mezclado con ligeros agregados de una herencia heterogénea. A esta mezcla en la que se nos aparece envuelto el tipo en cualquier caso particular la llamamos la *aleación constitucional.*

Las declaraciones metafóricas de este género son sugerentes, pero no bastan para la formulación de una teoría que explique de manera explícita y objetiva los casos impuros. Se necesita todo un arsenal conceptual para describir y distinguir aleaciones constitucionales en las que los tipos puros estén representados con diferentes intensidades. Por ejemplo, para dar un significado claro y objetivo a la noción de un tipo puro, digamos *A*, que diferentes individuos pueden representar en grados diferentes, se requieren criterios objetivos que permitan determinar, dados dos individuos cualesquiera, si representan el tipo *A* con igual intensidad, y de no ser así, cuál de ellos representa *A* con mayor intensidad que el otro. Los criterios adecuados de este género no efectuarán una división del universo del discurso en dos clases, *A* y no-*A*, sino un ordenamiento simple (casi lineal) del universo. En este ordenamiento dos individuos "coincidirán", es decir, ocuparán el mismo lugar si exhiben *A* con igual intensidad, en un sentido establecido por el criterio; mientras que el individuo *x* precederá al *y* si *x* es un caso de *A* menos pronunciado que *y*, en el sentido establecido por el criterio.

Un caso análogo tomado de la física servirá para ejemplificar este punto: puede efectuarse un ordenamiento simple de los minerales según su dureza mediante el criterio del rayado que ya hemos mencionado; si una punta de *y* raya una superficie de *x*, pero no a la inversa, *y* es más duro que *x* y, por lo tanto, sigue a *x* en el orden de dureza creciente; si *y* no es más duro que *x* ni éste más duro que *y*, ambos minerales ocuparán el mismo lugar en el orden casi lineal. Este ejemplo aclara dos puntos elementales pero importantes: 1) los criterios basados en "diagnósticos" que sirven para ubicar los casos individuales en el esquema no son criterios de pertenencia a una clase, como lo serían en un sistema estrictamente clasificatorio; más bien son criterios de precedencia y coincidencia en un orden casi lineal; 2) tales criterios pueden ser totalmente objetivos y bastante precisos sin presuponer mediciones cuantitativas.[20]

Hemos señalado que los sistemas tipológicos recientes han, en efecto, reemplazado un procedimiento estrictamente clasificatorio por un orde-

[19] Kretschmer (1925), pág. 93.

[20] Para un detallado análisis de los procedimientos ordenadores, con especial referencia a las teorías tipológicas, véase Hempel y Oppenheim (1936); en Hempel (1952), parte III, se encontrará una breve exposición general de la lógica de la clasificación, el ordenamiento y la medición.

namiento, (aunque algunos de ellos usen una terminología de tipo clasificatorio y la complementen hablando metafóricamente de casos límite, mezclas, formas de transición, etcétera). Ejemplifica tal adherencia a conceptos y métodos de carácter ordenador, no sólo el sistema de Kretschmer, sino también, para mencionar otros pocos ejemplos, la distinción de C. G. Jung entre casos extrovertidos e introvertidos, la tipología de E. R. Jaensch[21] y el sistema elaborado más recientemente por Sheldon en colaboración con Stevens y otros. Pero esta última teoría hace totalmente explícito el carácter ordenador de sus conceptos básicos y trata de satisfacer el requisito de objetividad (en el sentido examinado antes) para los criterios basados en diagnósticos que establece.

Puesto que cada uno de los tipos distinguidos en una teoría tipológica representará al menos un ordenamiento casi lineal, los sistemas tipológicos habitualmente realizarán un ordenamiento de los individuos a lo largo de diversos ejes, y de este modo reemplazarán los esquemas clasificatorios por "espacios" de referencia de varias "dimensiones".

Las ventajas del ordenamiento sobre la clasificación pueden ser considerables. En particular, el ordenamiento permite distinciones más sutiles que la clasificación. Además, puede adoptar la forma especial de un procedimiento cuantitativo, en el que cada dimensión está representada por una característica cuantitativa. Y los conceptos cuantitativos no sólo permiten una finura y precisión en las distinciones que no tienen paralelos en los niveles de la clasificación y del ordenamiento no cuantitativo, sino que también brindan una base para el uso de las poderosas herramientas de la matemática cuantitativa: es posible expresar leyes y teorías en términos de funciones que vinculen varias variables, y por medio de las técnicas matemáticas deducir de ellas consecuencias para los fines de la predicción y la prueba.

Las consideraciones presentadas en esta sección y en la precedente sugieren que la elaboración de conceptos taxonómicos en el estudio de los desórdenes mentales probablemente seguirá dos tendencias: primero, una continuación del paso de los sistemas definidos con referencia a características observables a sistemas basados en conceptos teóricos; y segundo, un paso gradual de los conceptos y métodos clasificatorios a los conceptos y procedimientos ordenadores, tanto del tipo cuantitativo como del no cuantitativo.

BIBLIOGRAFIA

American Psychiatric Association: *Diagnostic and Statistical Manual: Mental Disorders.* Washington, D. C., 1952.
Bridgman, P. W.: *The Logic of Modern Physics.* Nueva York, Macmillan, 1927.
Ellis, Albert: "An Operational Reformulation of Some of the Basic Principles of

[21] Véase, por ejemplo, Jung (1921) y Jaensch (1933).

Psychoanalysis". En Feigl, H. y Scriven, M. (comps.): *Minnesota Studies in the Philosophy of Science*, vol. I. Minneapolis, University of Minnesota Press, 1956, págs. 131-154.
Gregg, John R.: *The Language of Taxonomy*, Nueva York, Columbia University Press, 1954.
Hanson, N. R.: *Patterns of Discovery*. Londres, Cambridge University Press, 1958.
Hempel, Carl G.: *Fundamentals of Concept Formation in Empirical Science*. Chicago, University of Chicago Press, 1952.
Hempel, Carl G.: "The Theoretician's Dilemma". En Feigl, H., Scriven, M. y Maxwell G. (comps.): *Minnesota Studies in the Philosophy of Science*, vol. II. Minneapolis, University of Minnesota Press, 1958, págs. 37-98 (reproducido en este volumen).
Hempel, Carl G. y Oppenheim, P.: *Der Typusbegriff im Lichte der neuen Lokig*. Leiden, Sitjhoff, 1936.
Huxley, J.: *The New Systematics*. Oxford, Clarendon Press, 1940.
Jaensch, E. R.: *Die Eidetik und die typologische Forschungsmethode*. Leipzig, Quelle und Meyer, 1933.
Jevons, W. S.: *The Principles of Science*, 2ª ed., 1877; reimpreso con una nueva introducción de Ernest Nagel, Nueva York, Dover Publications, 1958.
Jung, C. G.: *Psychologische Typen*, Zurich, Rascher, 1921.
Kretschmer, E.: *Physique and Character*. Traducido de la segunda edición alemana por W. J. H. Sprott, Nueva York, Harcourt, Brace and Co., 1925.
Lazarsfeld P. y Barton, A. H.: "Qualitative Measurement in the Social Sciences: Classification, Typologies, and Indices". En Terner, D. y Lasswell, H. (comps.): *The Policy Sciences*. Stanford, Stanford Univ. Press, 1951, págs. 155-192.
Northrop, F. S. C.: *The Logic of the Sciences and the Humanities*, Nueva York, Macmillan, 1947.
Sheldon, W. H. y Stevens, S. S.: *The Varieties of Temperament*. Nueva York, Harper & Brothers, 1942.
Sheldon, W. H. y Tucker, W. B.: *The Varieties of Human Physique*. Nueva York, Harper & Brothers, 1940.
Simpson, George G.: *The Principles of Classification and a Classification of Mammals*. Boletín del American Museum of Natural History, vol. 45. Nueva York, 1945.
Simpson, George G., Pittendrigh, C. S. y Tiffany, L. H.: *Life: An Introduction to Biology*. Nueva York, Harcourt, Brace and Co., 1957.
Zubin, J. (comp.): *Field Studies in the Mental Disorders*. Nueva York, Grune and Stratton, 1961.

CAPÍTULO VII

Métodos tipológicos en las ciencias naturales y sociales *

1. *Introducción*

El concepto de tipo ha jugado un papel importante en distintas fases del desarrollo de la ciencia empírica. Muchos de sus usos tienen hoy día sólo interés histórico; pero algunas ramas de la investigación, especialmente la psicología y las ciencias sociales, han continuado empleando hasta el presente los conceptos tipológicos para fines descriptivos y teóricos. En particular, se han propuesto distintas tipologías caracterológicas y físicas, que proporcionan enfoques fructíferos al estudio de la personalidad. La investigación de tipos de constitución física y mental "extremos" o "puros" ha sido propugnada como un medio para comprender el funcionamiento de los individuos "normales". En cuanto a la ciencia social, el uso de tipos ideales se ha proclamado como una de las características metodológicas que la distinguen, en esencia, de la ciencia natural.

Tomando en cuenta estos usos recientes de los conceptos tipológicos y las distintas pretensiones concernientes a su peculiar significación podría ser interesante e importante a la vez lograr una clara comprensión de su status lógico y función metodológica. Existe una voluminosa bibliografía sobre el tema, pero que peca en gran parte de cierta falta de adecuación al aparato lógico utilizado para el análisis de los problemas planteados. En particular, muchos de los estudios que se dedican a la lógica de los conceptos tipológicos utilizan solamente los conceptos y principios de la lógica clásica, que en esencia es una lógica de propiedades o clases, y que, por lo tanto, no puede encarar adecuadamente las relaciones y los conceptos cuantitativos. Es ilustrativo de esta situación que Max Weber, quien con tanta elocuencia defiende el método de los tipos ideales en las ciencias sociales, emita un juicio claramente negativo sobre su status lógico: no pueden ser definidos por *genus proximum* y *differentia specifica,* y los casos concretos no pueden ser subordinados a ellos como instancias; [1] es decir, no son simplemente conceptos de clases o de

* Reproducido, con autorización del editor, del Acta del Simposio publicada bajo el auspicio de la American Philosophical Association (Asociación Filosófica Norteamericana), Eastern Division, con el título de: *Science, Language and Human Rights,* Filadelfia, University of Pennsylvania Press, 1952.

[1] Max Weber: *On the Methodology of the Social Sciences,* traducido y compi-

propiedad. En cambio, cuando se trata de una caracterización positiva, recurre a un lenguaje mucho menos preciso que con frecuencia es metafórico. De acuerdo con Weber, un tipo ideal es una construcción mental formada por una síntesis de muchos fenómenos individuales concretos que se hallan difusos, más o menos presentes y en ciertas ocasiones ausentes, ordenados de acuerdo con puntos de vista marcadamente unilaterales para conformar una estructura analítica unificada que por su pureza conceptual no puede encontrarse en la realidad: se trata de una utopía, un concepto limitador, con el cual sólo se pueden comparar los fenómenos concretos a fin de explicar algunos de sus componentes significativos.[2] Esta caracterización y muchos planteos similares que Weber y otros han hecho acerca de la naturaleza de los tipos ideales son ciertamente sugestivos, pero les falta claridad y rigor, por lo cual requieren un análisis lógico más profundo.

Además del status lógico de los conceptos tipológicos, me parece que habría que examinar de nuevo algunas de las demandas metodológicas que de ellos se pretende.

Este trabajo es, pues, un intento de explicar esquemáticamente el carácter lógico y metodológico de los conceptos tipológicos y evaluar su significación potencial para los propósitos a los que se pretende hacerlos servir.

El término "tipo" se ha utilizado en varios sentidos bastante diferentes entre sí. Propongo que se distingan aquí tres clases de conceptos de tipo que, como breve referencia y sujeto a una clarificación posterior, llamaremos tipos clasificatorios, extremos e ideales. Consideraré cada uno de ellos por separado.

2. *Tipos clasificatorios*

El uso de los conceptos de tipo para la clasificación está ilustrado por la teoría tipológica sobre el carácter y el físico de Ernst Kretschmer,[3] que ha tenido mucha influencia y en la cual los tipos se estructuran como clases. En este caso, la lógica del procedimiento tipológico es la bien conocida lógica de la clasificación, que no exige mayores explicaciones. Desde el punto de vista metodológico, la formación de tipos clasificatorios, igual que cualquier otro tipo de clasificación en la ciencia empírica, está sujeta a las exigencias de la utilidad sistemática. Las características que sirven para definir distintos tipos no sólo deben proporcionar casilleros perfectos para acomodar todos los casos particulares en el ámbito de la

lado por E. A. Shils y H. A. Finch. Nueva York, The Free Press of Glencoe, 1949, pág. 93.

[2] *Loc. cit.*, págs. 90-93.

[3] Ernst Kretschmer: *Physique and Character*, traducido por W. J. H. Sprott. Nueva York, Harcourt, Brace & World, 1936. Sobre teoría y técnica de la formación de tipos clasificatorios en la investigación social contemporánea, véase Paul F. Lazarsfeld y Allen H. Bartin, "Qualitative Measurement in the Social Sciences: Classification, Typologies, and Indices" en Daniel Lerner y Harold D. Lasswell (comps.): *The Policy Sciences*. Stanford, Stanford University Press, 1951.

investigación, sino que han de prestarse a generalizaciones certeras y proveer de esta manera una base para la predicción. Así, por ejemplo, las tipologías constitucionales a menudo apuntan a definir sus tipos, refiriéndolos a ciertas propiedades físicas que están asociadas empíricamente con una variedad de rasgos psicológicos, de modo que cada tipo represente un cúmulo de características concomitantes. Este objetivo es el núcleo metodológico de la búsqueda de clases o tipos "naturales", que se diferencia de aquella que lo intenta hacer "artificialmente".

En relación con los tipos clasificatorios, habría que hacer una breve referencia al uso del término "típico" cuando el mismo significa promedio, ya que ese uso evidentemente presupone una clasificación. Así, el enunciado de que un estudiante universitario norteamericano típico tiene, por ejemplo, 18,9 años, pretende afirmar el valor promedio de una cierta magnitud para una clase específica. Dado que existen distintos tipos de promedios y que ninguno proporciona mucha información sin una medida adicional de dispersión, es evidente que para cualquier propósito científico serio el uso del término "típico" ha de ser suplantado por una formulación más precisa en términos estadísticos.

3. *Tipos extremos*

Sin embargo, los intentos para llegar a una clasificación tipológica en la ciencia empírica se frustran a menudo frente a la percepción de que aquellas características de la temática que deberían proveer la base para definir la clasificación, no pueden construirse provechosamente como simples conceptos de propiedad que determinen, por extensión, clases con límites perfectamente demarcados. Así, si tratamos de formular criterios explícitos y precisos para distinguir entre personalidades extravertidas e introvertidas, pronto nos daremos cuenta de que la adopción de criterios de clasificación que marquen un límite preciso entre dos categorías se convierte en un procedimiento "artificial", teóricamente estéril. Parecería mucho más natural, o sea, sistemáticamente más promisorio, interpretar los dos conceptos como pasibles de gradaciones, de modo que un individuo dado no sea calificado como extravertido o introvertido sino como exhibiendo hasta cierto punto ambos rasgos. De esta manera las personalidades puramente extravertidas e introvertidas pasan a ser concebidas como tipos "extremos" o "puros", de los cuales rara vez o nunca se encontrarán instancias concretas, pero que podrán servir como puntos de referencia conceptuales o "polos" entre los cuales los casos particulares pueden ordenarse en una serie. En esta concepción general se basan varios sistemas de tipos psicológicos y físicos contemporáneos enunciados recientemente tales como la teoría de Sheldon sobre físico y temperamento.[4]

¿Cuál es la forma lógica de estos conceptos de tipos "extremos" o

[4] W. H. Sheldon, S. S. Stevens y W. B. Tucker: *The Varieties of Human Physique*. Nueva York y Londres, Harper and Row, 1940. También W. H. Sheldon y S. S. Stevens: *The Varieties of Temperament*. Nueva York y Londres, Harper and Row, 1942.

"puros"? Es evidente que no pueden construirse como conceptos de clase: los casos individuales no pueden subordinarse a ellos como instancias sino que sólo pueden caracterizarse en la medida en que se aproximen a ellos. En otras palabras, si el término *T* es un tipo extremo no puede decirse que un individuo *a* sea *T* o no-*T*; en cambio *a* podría ser, en cierto modo, "más o menos *T*". Ahora bien, ¿cómo deberá definirse objetivamente este "más o menos"? Una descripción de un tipo extremo con el cual puedan compararse casos concretos, por más vívida que sea, no puede por sí sola proveer estándares para tal comparación. En el mejor de los casos puede sugerir un programa de investigación que concentre la atención sobre ciertos fenómenos y regularidades empíricas, estimulando esfuerzos que ayuden a desarrollar un aparato conceptual preciso, adecuado para su descripción e interpretación teórica. Si un tipo extremo ha de funcionar como concepto científico legítimo en enunciados científicos que tengan significado objetivo claro, entonces deberán encontrarse criterios explícitos para comparar el "más o menos". Estos criterios pueden adoptar una forma no numérica, "puramente comparativa", o bien pueden basarse en mecanismos cuantitativos tales como las ponderaciones o mediciones.

La forma más simple, puramente comparativa, de un tipo extremo de concepto *T* puede especificarse sentando criterios que determinen para cualquiera de dos casos *a*, *b* en el universo que se investiga, si (i) *a* es más *T* que *b*, o (ii) *b* es más *T* que *a*, o (iii) *a* es tan *T* como *b*. Si tomamos el concepto de introversión pura como tipo extremo, por ejemplo, esto requeriría criterios objetivos que determinen para cualquiera de los dos individuos *a*, *b*, si son idénticamente introvertidos o si por el contrario uno de ellos es más introvertido que el otro. De esta manera se define un tipo extremo *T* de orden puramente comparativo u ordenador, no por el *genus* y *differentia* a la manera de un concepto de clase, sino especificando dos relaciones diádicas: "más *T* que" y "tan *T* como". Ahora bien, si el criterio que define esas relaciones ha de arrojar un ordenamiento de todos los casos particulares en una formación lineal que refleje el aumento de *T*, entonces debe responder a ciertas exigencias formales: "más *T* que" debe ser una relación asimétrica y transitiva; "tan *T* como" debe ser simétrica y transitiva y ambas deben satisfacer una ley de tricotomía en el sentido de que cualquier par de casos particulares *a*, *b* debe cumplir las condiciones de definición para exactamente una de las tres alternativas (i), (ii), (iii) mencionadas más arriba.[5]

La clase de concepto ordenador que aquí se caracteriza puede muy bien ilustrarse con una definición tomada de la mineralogía acerca del concepto puramente comparativo de dureza, refiriéndolo a prueba del rayado: Se dice que un mineral *a* es más duro que otro *b* si una punta filosa de una muestra de *a* raya la superficie de una muestra de *b*, pero no a la inversa. Si ninguno de los minerales es más duro, se dice que tienen la misma dureza. Así definidas, puede decirse que las dos relaciones determinan un tipo extremo, puramente comparativo, de dureza: pero esta

[5] Para más detalles véase Carl G. Hempel y Paul Oppenheim: *Der Typusbegriff im Lichte der neuen Logik.* Leyden, Holanda, Siythoff, 1936, cap. III.

terminología tendería a oscurecer más que a clarificar la lógica del procedimiento y por lo tanto no es utilizada.

En psicología y en las ciencias sociales es difícil, por decir lo menos, encontrar criterios objetivos útiles, análogos a aquellos basados en la prueba del rayado, que determinen un orden tipológico puramente comparativo. Por lo tanto encontramos que aquellos que proponen conceptos de tipo extremo, en tanto proporcionan criterios precisos y no caracterizaciones meramente sugestivas, terminan ya sea estructurando sus tipos como clases o bien especificando sus órdenes tipológicos con referencia a escalas de calsificación o procedimientos de medición que definen un "grado de *T*" numérico, por ejemplo. La primera línea se ejemplifica en la tipología de Kretschmer de físico y carácter: en ella usa el lenguaje de los tipos puros para una caracterización intuitiva del material que ha de investigar, mientras que para las formulaciones exactas erige a cada uno de los tipos principales como clase y acomoda los casos intermedios en algunas clases adicionales designadas como "tipos mixtos". El segundo curso está ejemplificado por la tipología del físico de Sheldon, que asigna a cada individuo una posición específica en cada una de las tres escalas de siete puntos que representan los rasgos de tipo básico de la teoría: endomorfismo, mesomorfismo y ectomorfismo.

Pero una vez que se han especificado criterios "operacionales" convenientes de orden estrictamente comparativo o cuantitativo, los tipos puros pierden su importancia especial. Ahora representan simplemente los puntos extremos de la gama, definida por los criterios dados; desde un punto de vista sistemático la terminología tipológica no significaría más que si se dijera que la conductividad eléctrica específica de un material dado indicase hasta qué punto se acerca al tipo extremo o puro de un conductor perfecto.

El uso de conceptos de tipo extremo como los que aquí hemos considerado, refleja un intento de avanzar desde el nivel clasificatorio y cualitativo de la formación del concepto, al cuantitativo: el proceso de ordenar conceptos de orden puramente comparativo representa una etapa lógica intermedia. En tanto falten criterios explícitos para su uso, como señaláramos antes, tendrán un status que en esencia es programático pero nunca sistemático. Una vez que se han especificado criterios convenientes, el lenguaje de los tipos extremos se hace innecesario porque no hay peculiaridades lógicas que diferencien a los conceptos de tipo extremo de otros conceptos comparativos y cuantitativos de la ciencia lógica: su lógica es la lógica del ordenamiento de las relaciones y de las mediciones. Por lo tanto nos referiremos a ellos también como *tipos ordenadores.*

Las tipologías de ordenamiento, al igual que las de clasificación, pertenecen por regla general a una etapa primitiva en el crecimiento de una disciplina científica, una etapa que se ocupa del desarrollo de un sistema de concepción en gran parte "empírico", que sirve para describir y para un bajo grado de generalización. La utilidad sistemática, que es una exigencia esencial en todas las etapas de la formación del concepto, consiste aquí, en el caso más simple, en una alta correlación entre criterios que "definen operacionalmente" un orden tipológico (tal como ciertos índices

antropométricos, por ejemplo) y una cantidad de otros rasgos variables (tales como otros índices anatómicos y fisiológicos o características psicológicas). Para las escalas cuantitativas tales correlaciones podrían tomar la forma, en casos favorables, de una proporcionalidad de diversas variables (análogas a la proporcionalidad, a temperatura constante, de las conductividades térmicas y eléctricas específicas de los metales) o bien pueden constituir otras relaciones invariantes, capaces de ser expresadas en términos de funciones matemáticas.[6]

4. *Tipos ideales y explicación en las ciencias sociales*

Como se mencionara en la primera sección, también los tipos ideales se presentan en general como resultado de aislar y exagerar ciertos aspectos de los fenómenos empíricos concretos, como conceptos limitativos que no pueden ejemplificarse totalmente y que en el mejor de los casos se aproximan a la realidad.[7] A pesar de la insinuación que encierra esta descripción creo que una reconstrucción lógica adecuada debe asignar a los tipos ideales un status diferente del de los tipos extremos o puros que describiera más arriba. Los tipos ideales o tipos construidos como Howard Becker los llama con toda propiedad, son introducidos en general sin ningún intento de especificar un orden de criterio apropiado y no se los utiliza para la clase de generalización característica de los tipos ordenadores; por el contrario, se los invoca como un mecanismo específico para la explicación de los fenómenos sociales e históricos. Trataré de argumentar ahora que esta concepción refleja un intento de promover la formación del concepto en sociología desde una etapa de descripción y

[6] Una explicación detallada sobre la lógica y metodología de los procedimientos cuantitativos y de ordenamiento puede encontrarse en Carl G. Hempel: *Fundamentals of Concept Formation in Empirical Science*, Chicago, University of Chicago Press, 1952, especialmente la sección 11. Acerca del uso de tales procedimientos en los estudios tipológicos, compárese con Lazarsfeld y Barton, obra citada; Hempel y Oppenheim, obra citada, y R. F. Winch: "Heuristic and Empirical Typologies: A Job for Factor Analysis", *American Sociological Review*, 12 (1947), 68-75.

[7] Una exposición detallada y versión crítica del concepto de tipo ideal, tal como se lo utiliza en la ciencia social puede encontrarse especialmente en las siguientes obras que han servido de guías al presente intento de análisis y reconstrucción:

Max Weber: *On the Methodology of the Social Sciences* (véase nota 1).

Max Weber: *The Theory of Social and Economic Organization*, traducido por A. M. Henderson y Talcott Parsons, Nueva York, Oxford University Press, 1947.

Alexander von Schelting: *Max Weber's Wissenschaftslehre*, Tübinger, J. C. B. Mohr, 1934.

Talcott Parsons: *The Structure of Social Action*, Nueva York, McGraw-Hill Book Company, 1937, cap. XVI.

Howard Becker: *Through Values to Social Interpretation*, Durham, N. C. Duke University Press, 1950.

Otras versiones críticas muy estimulantes sobre el concepto de tipo ideal pueden encontrarse en:

Félix Kaufmann: *Methodenlehre der Sozialwissenschaften*, Viena, Springer, 1936, especialmente la sección 6 de la segunda parte.

J. W. N. Watkins: "Ideal Types and Historical Explanation", *The British Journal for the Philosophy of Science*, 3 (1952), 22-43.

"generalización empírica", que está ejemplificada en la mayoría de los tipos clasificatorios y ordenadores, a la construcción de sistemas o módelos teóricos. A fin de ampliar y fundamentar esta opinión será necesario examinar con más detalle el carácter y función de los tipos ideales, tal como los concibieron sus propulsores.

De acuerdo con Weber y algunos autores que sostienen opiniones similares, el uso de tipos ideales hace posible explicar fenómenos sociales o históricos concretos tales como el sistema de castas en la India o el desarrollo del capitalismo moderno, en lo que tiene de individual y único. Se supone que esta comprensión consiste en aprehender las particularísimas relaciones causales que interconectan los elementos relevantes del hecho total que se está examinando. Si estas relaciones han de arrojar una explicación sociológicamente significativa, deberán ser no sólo "causalmente adecuadas" según este punto de vista, sino también tener sentido, es decir, que deben referirse a aspectos de la conducta humana que son impulsados de una manera inteligible por la evaluación u otros factores motivantes. Weber caracteriza los principios que expresan estas conexiones como "reglas empíricas generales" que se ocupan del modo en que los seres humanos tienden a reaccionar en situaciones dadas. Se dice que el "conocimiento nomológico" que transmiten se deriva de nuestra propia experiencia y de nuestro conocimiento de la conducta de los demás. Weber menciona la ley de Gresham como una generalización de esta especie: empíricamente bien fundamentada por la información pertinente de que se dispone, es "una interpretación racionalmente clara de la acción humana bajo ciertas condiciones y dado el presupuesto de que seguirá una línea puramente racional".[8]

En cuanto a modos específicos para descubrir principios explicativos significativos, Weber menciona el método de la comprensión empírica pero agrega la acotación de que no es universalmente aplicable ni tampoco siempre confiable. Por cierto, la experiencia subjetiva de la identificación empática con una figura histórica y de una comprensión inmediata (casi obviamente cierta) de sus motivaciones, no constituye un conocimiento, ni siquiera una comprensión científica, si bien puede ser una guía en la búsqueda de hipótesis generales explícitas del tipo requerido para una explicación sistemática. En realidad, la aparición de un estado empático en el intérprete no es condición ni necesaria ni suficiente para la correcta interpretación o comprensión en su sentido científico. No es necesaria, ya que una teoría apropiada de la conducta psicopática puede proporcionar al historiador una explicación de algunas fases de las acciones de Hitler aun faltándole la identificación empática. Ni es suficiente, ya que las hipótesis motivacionales sugeridas por la experiencia empática pueden ser objetivamente incorrectas.

Weber mismo señala que la verificación de la interpretación subjetiva siempre es indispensable; agrega que cuando no existen datos experimentales u observables adecuados "existe sólo el peligroso e incierto procedimiento del 'experimento imaginario', que consiste en desechar ciertos

[8] *The Theory of Social and Economic Organization*, pág. 98; compárese también con págs. 107-109.

elementos de la cadena de motivación y elaborar la línea de acción que probablemente surgiría, llegando así a un juicio causal".[9] Dado que sugiere qué habría pasado si ciertos componentes específicos de la situación hubiesen sido diferentes, el método arroja "juicios de posibilidad objetiva" que establecen la base de la imputación causal en las ciencias sociales. Estos juicios evidentemente presentan la forma de condicionales contrarios a los hechos y los investigadores de la lógica contrafáctica, tan discutida hoy en día, podrán interesarse por la fascinante ilustración de Weber del método propuesto con referencia a los problemas interpretativos de la historiografía, tales como la cuestión de la importancia de las Guerras Persas para el desarrollo de la cultura occidental;[10] la discusión de Weber de estos tópicos muestra con qué claridad este autor percibía la estrecha relación entre los condicionales contrarios a los hechos y las leyes generales.

Un tipo ideal, pues, debe servir como esquema de interpretación o explicación, incorporando un conjunto de "reglas empíricas generales" que establezcan conexiones "subjetivamente significativas" entre distintos aspectos de algún tipo de fenómeno, tal como la conducta económica perfectamente racional, una sociedad capitalista, la economía artesanal, una secta religiosa u otros. Entonces los tipos ideales no representan, por lo menos como intención, a los conceptos propiamente dichos, sino más bien a las teorías. Se insinúa pues la idea de que si esas teorías han de cumplir su propósito deben tener un carácter similar al de la teoría de los gases ideales, por ejemplo.[11] Para elaborar y fundamentar esta concepción trataré de demostrar primero que las supuestas diferencias entre el uso explicativo de los tipos ideales y el método de explicación en las ciencias naturales son espurias. Trataré a continuación de abordar un breve análisis comparativo del status de los conceptos "idealizados" y las teorías correspondientes en las ciencias naturales y sociales.

En ciencias naturales, explicar un hecho individual es explicar la presencia de alguna característica general o repetible (es decir, que pueda tener otras instancias; por ejemplo, un aumento de temperatura, la presencia de la corrosión, una caída de las presión arterial, etc.) en un caso particular, es decir, en un lugar específico o en un objeto determinado, en un momento dado o durante un cierto período (por ejemplo, el aire en New Haven durante las horas de la mañana del 5 de septiembre de 1952; el casco de un cierto barco; el paciente John Doe en un momento dado). La explicación de un hecho individual no puede significar razonablemente un relato de *todas* las características generales de una instancia particular

[9] *Loc. cit.*, pág. 97.

[10] *The Methodology of the Social Sciences*, págs. 164-188. Una esclarecedora amplificación y examen del análisis de Weber se pueden encontrar en von Schelting: obra citada, págs. 269-281.

[11] Naturalmente, se han realizado frecuentes paralelos entre tipos ideales y ciertas idealizaciones de la física (compárese por ejemplo con Weber: *The Theory of Social and Economic Organization*, pág. 110; Becker: obra citada, pág. 125). Sin embargo, es importante explicitar las similitudes involucradas y mostrar que no concuerdan con la pretensión de un status *sui generis* para conceptos de tipo ideal en las ciencias sociales.

dada, por ejemplo, *b*. Estas últimas implican el hecho de que en tales y tales direcciones y a tales y tales distancias espaciotemporales de *b* existen instancias particulares que tienen tales y tales propiedades generales. En consecuencia, explicar *todos* los aspectos generales de *b* es equivalente a explicar cada hecho individual del universo: pasado, presente y futuro. Es evidente que esta especie de explicación de un hecho particular "en su unicidad" es tan poco accesible para la sociología como para la física. En realidad hasta su *significado* preciso es bastante problemático. De esta manera, todo lo que puede buscarse coherentemente es la explicación de un hecho de alguna característica repetible *U* (que naturalmente puede ser bastante compleja) en una instancia particular dada *b*. La tarea de explicar el capitalismo occidental en su unicidad, por ejemplo, debe construirse así para ser de algún modo significativo; entonces sería análoga al problema de explicar el eclipse solar del 18 de marzo de 1950. En ambos casos existen ciertas características (cuya combinación se ha denominado *U* más arriba) para las que se busca una explicación (en el caso del eclipse estas características podrían incluir el dato de que el eclipse era anular, que no se vio en los Estados Unidos, que su duración fue de 5 horas y 42 minutos, etc.), pero existen otras innumerables características que no se toman en cuenta (tal como el número de diarios en los que se describe el hecho). Vale la pena señalar que el hecho que queremos explicar, *U(b)* para abreviar, sigue siendo único en el sentido de que la instancia particular *b* no puede repetirse. Si bien la existencia de otras instancias de *U* es posible, por lo menos desde el punto de vista lógico, ninguna de ellas podrá tener la misma localización espaciotemporal que *b*.

En las ciencias naturales un hecho particular se explica demostrando que su aparición puede inferirse por medio de leyes o principios teóricos de otras circunstancias particulares que en general los anteceden o son simultáneas. Lo que aclaran los escritos de Max Weber es que una explicación adecuada de un hecho cualquiera en sociología o historiografía tiene que tener, en esencia, el mismo carácter. La visión empática y la comprensión subjetiva no garantizan la validez objetiva, no constituyen una base para la predicción sistemática o la explicación de fenómenos específicos; estos últimos procedimientos deben basarse en principios empíricos generales, es decir, en el conocimiento nomológico. Por otra parte, las limitaciones que Weber propone a los principios explicativos de la sociología de reducirlos a reglas "significativas" de la conducta inteligible, son insostenibles. Muchos, por no decir todos, los hechos de interés para el científico social exigen que para explicarlos se haga referencia a factores "exentos de significado subjetivo" y por lo tanto a "uniformidades incomprensibles" para usar la terminología de Weber. Este autor reconoce que el sociólogo debe aceptar estos hechos como datos que tienen significado causal pero insiste en que "no alteran en lo mínimo la labor específica del análisis sociológico..., que es la interpretación de la acción en términos de su significado subjetivo".[12] Pero esta concepción excluye del campo de la sociología cualquier teoría de la conducta que renuncie al uso de

[12] *The Theory of Social and Economic Organization*, pág. 94.

conceptos motivacionales "subjetivamente significativos". Esto quiere decir, ya sea una restricción arbitraria del concepto de sociología —el que, como resultado, puede eventualmente llegar a ser inaplicable a cualquier otra rama de la investigación científica— o bien implicaría un juicio *a priori* en cuanto al carácter de cualquier sistema de conceptos que pueda de algún modo producir una teoría sociológica explicativa. Es evidente que este veredicto *a priori* no puede justificarse y que los recientes progresos en la teoría psicológica y social indican que es posible formular principios explicativos para la acción intencional en términos puramente conductistas y no introspectivos.

A continuación abordaremos el rol jugado por los experimentos imaginarios que, por otra parte, son bien conocidos en las ciencias naturales, y en los cuales será útil distinguir *dos tipos de experimentos imaginarios: los intuitivos y los teóricos.* Un experimento imaginario intuitivo apunta a anticipar el resultado de un procedimiento experimental puramente imaginado pero que puede muy bien ser realizado. La predicción está guiada aquí por la experiencia pasada con respecto a fenómenos particulares y sus regularidades, como también, ocasionalmente, por la creencia en ciertos principios generales que se aceptan como si fuesen verdades apriorísticas. De esta manera, al explicar la pareja distribución de resultados que se obtiene al tirar un dado regular o al anticipar resultados similares para un juego con un dodecaedro homogéneo regular, se invocan a menudo ciertas reglas de simetría, tales como el principio de la razón insuficiente. A veces se aducen principios similares en experimentos imaginarios con palancas y otros sistemas físicos que presentan ciertos rasgos de simetría. Los experimentos imaginarios de esta clase son intuitivos en el sentido de que los supuestos y datos en que se basa la predicción no se explicitan y ni siquiera entran en el proceso consciente de anticipación: la experiencia pasada y la creencia —posiblemente inconsciente— en ciertos principios generales funcionan aquí como guías sugestivas de la anticipación imaginativa, más que como base teórica para una predicción sistemática.

Por otra parte, el experimento imaginario de tipo teórico presupone un conjunto de principios generales explícitamente enunciados —tales como leyes de la naturaleza— y anticipa el resultado del experimento por inferencia deductiva o probabilística de aquellos principios, combinados con condiciones limitativas adecuadas que representen aspectos relevantes de la situación experimental imaginada. A veces, esto último no puede realizarse en la práctica, como cuando se deducen las leyes para un péndulo matemático ideal o para el impacto elástico perfecto a partir de principios más generales de la mecánica teórica. La pregunta de qué *pasaría si,* digamos, el hilo del péndulo fuese infinitamente fino y perfectamente rígido y *si* la masa del péndulo se concentrara en el extremo libre del hilo, se contesta aquí, no "desechando" aquellos aspectos de un péndulo físico que se desvían de este supuesto y tratando luego de dilucidar el resultado, sino por la reducción rigurosa a partir de principios teóricos existentes. Aquí no entra la imaginación: el experimento sólo es imaginario en el sentido de que la situación a la cual se refiere no se produce en la realidad y de que además podría ser técnicamente imposible ejecutarlo.

Los dos tipos de experimentos imaginarios aquí diferenciados constituyen tipos extremos, por decirlo así, que raramente se realizan en su forma pura: en muchos casos, los supuestos empíricos y el razonamiento subyacente en un experimento imaginario se detallan sobremanera aunque no son totalmente explícitos. Los diálogos de Galileo contienen excelentes ejemplos de este procedimiento que muestra hasta qué punto el método puede ser fructífero para sugerir una percepción teórica general. Por supuesto que los experimentos imaginarios intuitivos no sustituyen la recolección de datos empíricos mediante la experimentación o procedimientos de observación. Una buena ejemplificación de ello son los numerosos experimentos imaginarios, intuitivamente bastante plausibles, que se han aducido a fin de refutar la teoría especial de la relatividad. En cuanto a la experimentación imaginaria en las ciencias sociales, su resultado tiende a estar afectado por ideas preconcebidas, estereotipos y otros factores perturbantes. En la revisión del libro *The American Soldier*, de Stouffer, Lazarsfeld [13] hace un listado de una serie de supuestos psicológicos y sociológicos que parecen ser tan obviamente ciertos como para no necesitar más investigación, pero que en realidad no reciben ninguna confirmación en los resultados del grupo de Stouffer. Por ejemplo, los siguientes: que entre los soldados norteamericanos durante la guerra, los hombres de mayor nivel educacional mostraban más síntomas psiconeuróticos que los menos instruidos; que los soldados del Sur de los Estados Unidos podían aguantar mejor que los del Norte el clima caluroso de los Mares del Sur, etc. Las creencias de este tipo podían evidentemente afectar el resultado y derrotar el propósito de los experimentos intuitivos en sociología. Por lo tanto, estos experimentos no pueden proporcionar una evidencia pertinente para probar las hipótesis sociológicas. En el mejor de los casos pueden cumplir una función eurística: la de *sugerir* hipótesis que, sin embargo, deban someterse luego a pruebas objetivas apropiadas.

Los experimentos imaginarios mencionados por autores tales como Max Weber y Howard Becker como método de investigación sociológica son, evidentemente, de carácter intuitivo: su función eurística es ayudar a descubrir conexiones regulares entre varios componentes de la misma estructura o proceso social. Estas conexiones pueden luego incorporarse a un tipo ideal y proporcionar de este modo la base para usarlo como explicación de este último.

5. *Tipos ideales y modelos teóricos*

Hemos sostenido que, dado que los tipos ideales se proponen proveer de explicaciones, deben construirse como sistemas teóricos que incorporen hipótesis generales verificables. ¿Hasta qué punto se puede reconciliar esta concepción con la frecuente insistencia por parte de los propugnadores del método, de que los tipos ideales no pretenden ser hipótesis verificables con la evidencia empírica, de que la desviación del hecho concreto pertene-

[13] *Public Opinion Quarterly*, 13 (1949), págs. 377-404.

ce a su esencia misma? Consideremos más detalladamente cómo conciben la aplicación de los conceptos del tipo ideal a fenómenos concretos aquellos que defienden este punto de vista. Existen pocos enunciados precisos sobre el tema; tal vez la formulación más explícita es la que da Howard Becker cuando intenta desarrollar lo que denomina "una fórmula lógica para la tipología". Becker sostiene que los tipos ideales o construidos funcionan en las hipótesis que tienen la forma "Si *P* entonces *Q*", donde *P* es el tipo invocado y *Q* es una característica más o menos compleja.[14] En cuanto a la aplicación de tales hipótesis a los datos empíricos, dice Becker: "Sin embargo, en la naturaleza misma de la construcción de tipos, rara vez o nunca el consecuente se desprende empíricamente y el antecedente es, pues, empíricamente 'falso'. Si *Q'* entonces *P'* ".[15] Por medio de esta desviación del hecho empírico, ya que ocurre *Q'* y no *Q*, un tipo construido adquiere lo que Becker denomina "utilidad negativa"; promueve una búsqueda de factores distintos de los representados por *P* para explicar la discrepancia.[16] De esta manera, según Becker, "la tipología constructiva hace uso *planificado* de la estipulación 'siendo iguales o irrelevantes todas las demás condiciones', con el fin de determinar la 'desigualdad' o 'relevancia' de las 'otras condiciones' ".[17]

Esta opinión merece un análisis más detallado ya que sugiere —tal vez inintencionalmente— el uso de la cláusula *ceteris paribus* para una defensa convencionalista de las hipótesis tipológicas contra cualquier prueba concebible que no la confirme.[18] Para ilustrar este punto imaginemos a un físico que proponga la hipótesis de que en condiciones ideales, es decir, en un vacío cercano a la superficie de la Tierra, un cuerpo con caída libre durante t segundos cubrirá una distancia de exactamente $5t^2$ metros. Supóngase ahora que un experimento cuidadoso arroje resultados distintos de los que exige la hipótesis. Es evidente entonces que el físico no puede contentarse simplemente con inferir que las condiciones ideales exigidas no fueron cumplidas: además de esta posibilidad, debe admitir la alternativa de que la hipótesis que se está probando no es correcta. Para formular este punto en términos del esquema general de Becker diremos que se podría inferir que *P* no se cumple sólo si, además del resultado de la observación *Q'* pudiéramos dar por sentada la corrección de la hipótesis

[14] Obra citada, págs. 259-264. Becker describe la conexión entre *P* y *Q* como "probabilidad objetiva". Pero como usa la expresión 'Si *P* entonces *Q*' en una inferencia de la forma *modus tollens*, que es válida para la implicancia probabilística (es decir, para los enunciados que tienen la forma 'Si *P* entonces probablemente *Q*') parecería más adecuado reconstruir el comentario de Becker como si significase que 'Si *P* entonces *Q*' es una hipótesis que expresa una generalización empírica en el dado por Weber. Tal generalización, igual que cualquier otra hipótesis empírica, puede naturalmente sólo ser probable y nunca segura, con respecto a cualquier cuerpo de evidencia fáctica pertinente.

[15] Obra citada, pág. 262.

[16] Max Weber ha señalado de un modo similar la utilidad heurística de los tipos ideales; compárese por ejemplo *The Methodology of the Social Sciences*, págs. 90, 101-103; *The Theory of Social and Economic Organization*, pág. 111.

[17] Howard Becker: obra citada, pág. 264.

[18] Sobre la utilización de la cláusula *ceteris paribus*, véase también el excelente trabajo de Félix Kaufmann, *Methodology of the Social Sciences*, Nueva York, Oxford University Press, 1944, pág 84 y siguientes; pág. 213 y siguientes.

'Si *P* entonces *Q*'. Pero para este supuesto realmente no tenemos garantía: si la tuviéramos, la prueba en su totalidad no tendría sentido. Así, cuando sucede *Q'*, podemos tan sólo inferir ya sea que *P* no ocurrió o que la hipótesis 'Si *P* entonces *Q'* es falsa.

Parecería pues que se podría defender con convicción nuestra hipótesis tipológica si se la calificara con una cláusula *ceteris paribus* apropiada y darle así la forma: "Siendo iguales o irrelevantes todos los demás factores, *Q* siempre ocurrirá cuando ocurre *P*". Es evidente que ninguna prueba empírica podrá descalificar jamás una hipótesis que tenga esta forma ya que un resultado aparentemente desfavorable podrá siempre atribuirse a una violación de la cláusula *ceteris paribus* por interferencia de otros factores que son los específicamente incluidos en *P*. En otras palabras, la hipótesis ponderada puede no presentar excepciones gracias a la convención de invocar una violación de la cláusula *ceteris paribus* cada vez que un hecho de *P* no esté acompañado por un hecho de *Q*. Pero la misma convención que hace irrefutable a la hipótesis, también la exime de todo contenido empírico y por lo tanto de poder explicativo. Como la cláusula protectora no especifica *qué* factores distintos de *P* deben ser iguales (es decir, constantes) o irrelevantes si ha de garantizarse la predicción de *Q*, la hipótesis no tiene aplicación predictiva a los fenómenos concretos. De un modo similar, la idea de verificar una hipótesis dada pierde sentido. Es interesante señalar aquí por contraste, que en la formulación de hipótesis físicas nunca se usa la cláusula *ceteris paribus*: todos los factores que se consideran relevantes deben enunciarse explícitamente (como en la ley de la gravedad de Newton o en las leyes de Maxwell) o ser perfectamente comprensibles (como en la conocida formulación de la ley de Galileo que se refiere a la caída libre en un vacío cercano a la superficie de la Tierra); todos los demás factores, por implicación, se consideran irrelevantes. La prueba empírica es por lo tanto significativa y el descubrimiento de evidencias que no concuerdan exige revisiones apropiadas, ya sea modificando las supuestas conexiones funcionales entre variables que se han individualizado como relevantes, o bien introduciendo nuevas variables relevantes. La hipótesis de tipos ideales deberán ajustarse a las mismas normas para poder ofrecer una explicación teórica a los fenómenos históricos y sociales en lugar de ser esquematismos conceptuales vacíos de todo contenido empírico.

Sin embargo, ¿no es cierto que existen también en la física teorías tales como la de los gases ideales, la del impacto perfectamente elástico, la del péndulo matemático, la de los aspectos estadísticos del juego con un dado perfecto, etc., que no son invalidadas por el hecho de no poseer una ejemplificación precisa en el mundo empírico? ¿No podrían en ese caso exigir los tipos ideales, el mismo status que los conceptos centrales de esas teorías "idealizadas"? Estos conceptos se refieren a sistemas físicos que se satisfacen en ciertas condiciones extremas y que no pueden llenarse totalmente sino sólo por aproximación, a los fenómenos empíricos concretos. Su importancia científica reside, creo, en los siguientes puntos: a) las leyes que gobiernan la conducta de los sistemas físicos ideales pueden deducirse de principios teóricos más globales que están perfectamente con-

firmados por la evidencia empírica; la reducción en general adopta la forma de asignar ciertos valores extremos a algunos de los parámetros de la teoría global. Así, por ejemplo, las leyes de un gas ideal se obtienen a partir de principios más inclusivos de la teoría cinética de los gases "suponiendo" que los volúmenes de las moléculas de gas desaparecen y que no existen fuerzas de atracción entre las moléculas, es decir, igualando los parámetros apropiados a cero. b) Las condiciones extremas que caracterizan el caso ideal pueden aproximarse por lo menos empíricamente; y cada vez que el mismo se trata en una instancia concreta, las leyes ideales en cuestión quedan confirmadas empíricamente. Así, por ejemplo, la ley de los gases ideales de Boyle-Charles es satisfecha con bastante aproximación por una gran variedad de gases dentro de amplias gamas especificables de presión y temperatura (para una masa fija de gas) y por esta misma razón la ley puede invocarse de un modo significativo con fines explicativos.

El análisis precedente sugiere las siguientes observaciones con respecto a los aspectos ideales y empíricos de los conceptos de tipo ideal en las ciencias sociales:

(i) Las construcciones ideales no tienen el carácter de conceptos en su sentido más restrictivo, sino de sistemas teóricos. La introducción de tal estructura dentro de un contexto teórico requiere, entonces, no tanto la definición por género y diferencia, sino la especificación de un conjunto de características (tales como la presión, temperatura y volumen en el caso del gas ideal) *y también* un conjunto de hipótesis generales que relacione dichas características.

(ii) Por lo tanto, un concepto idealizado *P*, *no* funciona en hipótesis que tengan la forma simple 'Si *P* entonces *Q*'. Así, por ejemplo, la hipótesis 'Si una sustancia es un gas ideal entonces cumple la ley de Boyle' que tiene esa forma, es un enunciado analítico transmitido por la definición de gas ideal; no sirve a los fines de una explicación. Por el contrario, las hipótesis que caracterizan el concepto de gas ideal relacionan ciertas características cuantitativas de un gas y cuando se las aplica a sistemas físicos concretos, permiten predicciones empíricas específicas. Entonces, para decirlo de un modo tal vez demasiado simplificado, lo que entra en la teoría física no es de ningún modo el concepto de gas ideal, sino más bien los conceptos que representan las distintas características que se manejan en la teoría de los gases ideales: sólo que se las menciona en los principios de la termodinámica.

(iii) En las ciencias naturales, por lo menos, se considera que un conjunto de hipótesis caracteriza un sistema ideal solamente si representa lo que podría llamarse idealizaciones más bien *teóricas* que *intuitivas*, es decir, si se las puede obtener dentro del marco de referencia de una teoría dada, como casos especiales de principios más incluyentes. Así, por ejemplo, la fórmula para el péndulo matemático, descubierta empíricamente por Galileo, no constituyó una idealización teórica hasta después del establecimiento de hipótesis más globales que a) tuviesen confirmación empírica

independiente, b) incorporasen la fórmula del péndulo como caso especial, c) nos permitieran juzgar el grado de idealización involucrado en este último por medio de una descripción de factores adicionales relevantes al movimiento de péndulo físico y cuya influencia es bastante pequeña en el caso de aquellos sistemas físicos a los cuales usualmente se aplica esa fórmula.

Por supuesto, ninguna teoría, por más incluyente que sea, puede pretender dar una descripción completamente exacta de cualquier clase de fenómeno empírico. Siempre es posible que hasta la teoría más global y mejor confirmada sea perfeccionada en el futuro por la inclusión de nuevos parámetros y leyes apropiadas: la teoría más integral de hoy en día puede no ser mañana más que una idealización sistemática dentro de un marco teórico más amplio.

Entre los conceptos de tipos ideales en la teoría social, aquellos que se utilizan en la economía analítica son los que más se aproximan al status de idealizaciones en la ciencia natural: los conceptos de libre competencia perfecta, de monopolio, de conducta económica racional por parte de un individuo o de una empresa, etc., representan todos esquemas para la interpretación de ciertos aspectos de la conducta humana e involucran el supuesto idealizante de que los factores no económicos, del tipo de los que de hecho influyen sobre los actos humanos, pueden dejarse de lado para el propósito que nos interesa. En el contexto de la construcción rigurosa de una teoría, estas estructuras ideales reciben un significado preciso en forma de hipótesis que postulan conexiones matemáticas específicas entre ciertas variables económicas. Es frecuente que dichos postulados caractericen el tipo ideal o de conducta como aquel que maximiza una función dada de esas variables (por ejemplo, el beneficio).

Sin embargo, me parece que en dos aspectos importantes las idealizaciones en economía difieren de las de las ciencias naturales, principalmente por ser más bien idealizaciones intuitivas que teóricas, en el sentido de que los postulados correspondientes no se deducen, como casos especiales, de una teoría más amplia que cubra también los factores no racionales y no económicos que afectan la conducta humana. No existe por el momento una teoría adecuada más general y por lo tanto no hay base teórica para una apreciación de la idealización que se maneja al aplicar las construcciones económicas a situaciones concretas. Esto nos lleva a un segundo punto de la diferencia: la clase de fenómenos concretos de conducta para los cuales los principios idealizados de la teoría económica deben constituir, por lo menos aproximadamente, generalizaciones correctas, no siempre quedan claramente especificados. Esto obstaculiza, por supuesto, la validez de su uso como explicación de esos principios. Un sistema teórico ideal, igual que cualquier sistema teórico, sólo puede alcanzar el estado de mecanismo explicativo y predictivo si se ha especificado su área de aplicación. En otras palabras, si sus conceptos constitutivos han recibido interpretación empírica que, de manera directa o por lo menos mediata, los liga a fenómenos observables. Así, por ejemplo, el área de aplicación para la teoría de los gases ideales podría indicarse, en términos generales, interpretando los parámetros teóricos 'P', 'V', 'T' en

términos de las magnitudes "operacionalmente definidas" de presión, volumen y temperatura de los gases a presiones moderadas o bajas y a temperaturas moderadas o altas. De modo similar, la aplicación empírica de los principios de un sistema económico ideal requiere una interpretación en términos empíricos que no hacen analíticos a aquellos principios. Por lo tanto, la interpretación no debe atenerse al enunciado de que las proposiciones de la teoría son válidas para todos los casos de conducta económica racional, lo que equivaldría tan sólo a una tautología. En cambio, debe caracterizar, por medio de criterios lógicamente independientes de la teoría, aquellos tipos de conducta individual o grupal a los cuales pretende aplicarse la teoría. Con respecto a los mismos, deberá entonces anexarse una interpretación operacional bastante definida a parámetros teóricos, tales como 'dinero', 'precio', 'costo', 'beneficio', 'utilidad', etcétera. De este modo, las proposiciones de la teoría adquieren valor empírico: serán susceptibles de verificación y por lo tanto aun de no ser confirmadas. He aquí una característica esencial de todos los sistemas explicativos potenciales.

Los resultados de la comparación precedentes entre construcciones ideales de la economía y de la física no deben, sin embargo, considerarse como indicadores de una diferencia metodológica esencial entre ambas disciplinas. Haciendo referencia al primero de los dos puntos de la comparación, sólo deberá recordarse que gran parte del esfuerzo dirigido actualmente a la teorización en sociología tiende hacia el desarrollo de una teoría integral de la economía, en tanto permitan una aplicación empírica y puedan entonces alcanzar el status de idealizaciones más bien teóricas que intuitivas. Dejando de lado el logro de esa meta ambiciosa, es evidente que se requiere una interpretación para cualquier sistema teórico que tenga validez empírica: tanto en las ciencias sociales como en las naturales.

Los tipos ideales que se invocan en otros campos de la ciencia social carecen de la claridad y precisión presentes en las construcciones que utiliza la economía teórica. Las regularidades de conducta que intentan definir un tipo ideal dado, se enuncian en general en términos más o menos intuitivos y los parámetros que supuestamente deben relacionar, no se especifican explícitamente. Por último, no hay una indicación clara del campo de su aplicación empírica y la consecuente verificación que se pretende para el sistema tipológico. De hecho la exigencia de tal verificación a menudo se rechaza de un modo integral que, como lo demuestra la discusión precedente, creo, es poco consistente con la pretensión de que los tipos ideales proporcionen una comprensión de ciertos fenómenos empíricos.

Si el análisis esbozado aquí es en esencia correcto, entonces los tipos ideales podrán con toda seguridad servir a su propósito sólo si se los introduce como sistemas teóricos interpretados, es decir: a) especificando una lista de características de las cuales deberá ocuparse la teoría; b) formulando un conjunto de hipótesis en función de dichas características; c) dando esas características una interpretación empírica que asigne a la teoría un ámbito específico de aplicación; y d) como objetivo de largo alcance, incorporando el sistema teórico como caso especial, dentro de una

teoría más global. Hasta qué punto pueden lograrse estos objetivos no puede decidirse por medio del análisis lógico. Sin embargo, sería engañoso creer que cualquier procedimiento conceptual que en esencia carezca de las tres primeras consideraciones pueda brindar una comprensión teórica en cualquier campo de la investigación científica. En la medida en que el programa aquí esbozado puede llevarse a la práctica, el uso de tipos ideales es más bien un aspecto terminológico sin importancia que una característica metodológica diferenciada de las ciencias sociales: el método de los tipos ideales ya no se distingue de los métodos utilizados por otras disciplinas científicas en la formación y aplicación de conceptos y teorías explicativos.

6. *Conclusión*

En resumen: los distintos usos de los conceptos de tipo en psicología y en las ciencias sociales, cuando se los libera de ciertas connotaciones que pueden conducir a error, exhiben básicamente el mismo carácter que los métodos de clasificación, ordenamiento, medición, correlación empírica y formación de teoría utilizados en las ciencias naturales. Dirigido hacia este resultado, el análisis de procedimientos tipológicos muestra una importante similitud lógica entre diversas ramas de la ciencia empírica.

CAPÍTULO VIII

El dilema del teórico: un estudio sobre la lógica de la construcción de teorías [1]

1. *Sistematización deductiva e inductiva*

La investigación científica en sus diversas ramas no busca simplemente registrar sucesos particulares en el mundo de nuestra experiencia, sino que trata de descubrir regularidades en el flujo de los acontecimientos y formular así leyes generales que puedan usarse para la predicción, la retrodicción [2] y la explicación.

Los principios de la mecánica de Newton, por ejemplo, hacen posible, dadas las posiciones y los impulsos actuales de los objetos celestes que componen el sistema solar, predecir sus posiciones e impulsos para un instante determinado del futuro o retrodecirlos para una instante determinado del pasado; de modo similar, esos principios permiten explicar las posiciones y los impulsos actuales sobre la base de los que tenían en un instante anterior. Además de dar cuenta así de hechos particulares como los que se acaba de mencionar, los principios de la mecánica de Newton también explican ciertos "hechos generales", o sea, uniformidades

[1] Este artículo ha sido reimpreso, con algunos cambios, con permiso del editor. Extraído de *Minnesota Studies in the Philosophy of Science*, vol. II, editado por Herbert Feigl, Michael Scriven y Grover Maxwell, University of Minnesota Press, Minneapolis, Copyright 1958 por la Universidad de Minnesota.

[2] Este término * fue sugerido por un pasaje del libro de Reichenbach (1944), en el cual se aplica la palabra 'retrodictibilidad' a la posibilidad de determinar "datos del pasado en términos de observaciones dadas" (pág. 13). En un contexto similar, Ryle usa el término 'retrodecir' (véase, por ejemplo, 1949, pág. 124); y Walsh habla de la tarea de los historiadores "de 'retrodecir' el pasado: establecer sobre la base de elementos de prueba actuales cómo debe haber sido el pasado" (93, páginas 41). De acuerdo con una observación que aparece en el análisis del libro de Walsh que hace Acton (*Mind*, volumen 62 [1953], págs. 564-565), la palabra 'retrodicción' ya fue usada en el mismo sentido por J. M. Robertson en *Buckle and his Critics* (1895).

* Las palabras inglesas correspondientes a 'retrodicción' y 'retrodictibilidad' utilizadas por Reichenbach son '*postdiction*' y '*postdictability*' mientras que las correspondientes a 'retrodecir' y 'retrodicción' utilizadas por Ryle, Walsh y Robertson son '*retrodict*' y '*retrodiction*'. [T.]

empíricas como las leyes del movimiento planetario de Kepler ya que éstas pueden deducirse de aquéllos.[3]

La explicación, la predicción y la retrodicción científicas tienen el mismo carácter lógico: muestran que se puede inferir el hecho en cuestión a partir de otros hechos por medio de determinadas leyes generales. En el caso más simple, este tipo de argumento puede esquematizarse como una inferencia deductiva de la forma siguiente:

$$(1.1) \qquad \frac{C_1, C_2 \ldots C_k \quad L_1, L_2 \ldots L_r}{E}$$

Aquí, $C_1, C_2, \ldots C_k$ son enunciados de sucesos particulares (por ejemplo, de las posiciones e impulsos de ciertos cuerpos celestes en un determinado instante), y $L_1, L_2, \ldots, L_r$, leyes generales (por ejemplo, las de la mecánica de Newton); finalmente, E es una proposición que enuncia aquello que se explica, predice o retrodice. El argumento sólo tiene fuerza si su conclusión, E, se sigue deductivamente de las premisas.[4]

Mientras que la explicación, la predicción y la retrodicción son semejantes en su estructura lógica, difieren en otros aspectos. Por ejemplo, un argumento de la forma (1.1) será calificado como una predicción sólo si E se refiere a un suceso posterior al instante en que se formula el argumento; en el caso de una retrodicción, el acontecimiento debe ocurrir antes de la presentación del argumento. Sin embargo, esas diferencias no requieren aquí un estudio más completo ya que el propósito de la exposición precedente fue simplemente señalar el papel que desempeñan las leyes generales en la explicación, la predicción y la retrodicción científicas.

Para esos tres tipos de procedimiento científico, usaremos el término común '*sistematización* (*deductiva*)'. Es más preciso decir que el término se usará para hacer referencia, primero, a cualquier argumento de la forma (1.1) que satisfaga los requisitos indicados antes, ya sea que sirva como explicación, predicción o retrodicción o aun con otro carácter; segundo, al procedimiento para formular argumentos del tipo que acabamos de caracterizar.

Hasta aquí hemos considerado sólo aquellos casos de explicación, predicción y procedimientos afines que pueden considerarse como argumentos deductivos. Hay muchos casos de explicación y de predicción científica, sin embargo, que no encajan en una pauta estrictamente deductiva. Se puede explicar, por ejemplo, que Juan ha contraído el sarampión, señalando que se contagió de su hermana, convaleciente de esa enfermedad. Los hechos particulares antecedentes que se aducen aquí son que

[3] Más exactamente: de los principios de la mecánica de Newton puede deducirse que las leyes de Kepler valen con aproximación si se supone que las fuerzas ejercidas sobre los planetas por los objetos celestes diferentes del sol (especialmente otros planetas) son despreciables.

[4] Para una presentación más completa de este esquema y para ciertas salvedades concernientes a la identidad estructural de los argumentos explicativo y predictivo, véase el capítulo XII.

Juan se expuso al contagio y, supongámoslo, que no había tenido antes el sarampión. Pero, para relacionarlos con el acontecimiento a explicar, no podemos aducir una ley general según la cual en las circunstancias especificadas el sarampión se transmite invariablemente a la persona expuesta al contagio: sólo se puede afirmar que hay una gran probabilidad de transmisión (en el sentido de frecuencia estadística). Puede usarse también el mismo tipo de argumento para la predicción o la retrodicción de un caso de sarampión.

De modo similar, en una explicación psicoanalítica de la conducta neurótica de un adulto realizada sobre la base de ciertas experiencias de la infancia, las generalizaciones que pueden aducirse para relacionar los acontecimientos-antecedentes con aquellos a explicar, en el mejor de los casos se podrán considerar como enunciados que establecen probabilidades más o menos elevadas para las conexiones en cuestión, pero no como expresiones de uniformidades sin excepciones.

Las explicaciones, predicciones y retrodicciones del tipo que se ha ejemplificado aquí difieren de las expuestas previamente en dos aspectos importantes: las leyes aducidas son de forma diferente y el enunciado que se ha de formular no se sigue deductivamente de los enunciados explicativos presentados. Examinaremos ahora esas diferencias un poco más de cerca.

Las leyes a que se ha hecho referencia en relación con el esquema (1.1), tales como las de la mecánica de Newton, son *enunciados de forma estrictamente universal* o *enunciados estrictamente universales*. Un enunciado de este tipo es una afirmación, que puede ser verdadera o falsa, según la cual todos los casos que satisfacen determinadas condiciones tendrán sin excepción tales y cuales características. Por ejemplo, el enunciado 'Todos los cuervos son negros' es una oración de forma estrictamente universal; y lo es también la primera ley del movimiento de Newton según la cual todo cuerpo persiste en su estado de reposo o de movimiento rectilíneo a una velocidad constante.

Las leyes invocadas en el segundo tipo de argumentos explicativos y de otros relacionados con éstos, son, por otra parte, de *forma estadística;* son *enunciados de probalilidad estadística*. Un enunciado de este tipo es la afirmación —que puede ser verdadera o falsa— que establece, para los casos que satisfacen determinadas condiciones, la probabilidad de que tengan tales y cuales características.[5]

[5] La distinción que se ha hecho aquí se refiere, pues, exclusivamente a la *forma* de los enunciados en consideración y no a su valor de verdad ni a la medida en que están apoyados por elementos de prueba empíricos. Si se estableciera, por ejemplo, que en realidad sólo el 80 % de todos los cuervos son negros, esto no mostraría que 'todos los cuervos son negros' (o E_1 para abreviar) es un enunciado de probabilidad estadística sino que es un enunciado falso de forma estrictamente universal, y que 'la probabilidad de que un cuervo sea negro es de 0,80' (o E_2 para abreviar) es un enunciado verdadero de forma estadística.

Más aún, con seguridad ni E_1 ni E_2 pueden fundamentarse jamás de modo concluyente: sólo pueden estar más o menos apoyados por elementos de prueba disponibles; cada uno de ellos tiene así una probabilidad lógica o inductiva, más o menos elevada en relación con esos elementos de prueba. Pero esto a su vez no afecta para nada el hecho de que E_1 sea de forma estrictamente universal y E_2, de forma estadística.

Para distinguirlos con pocas palabras: un enunciado estrictamente universal del tipo más simple tiene esta forma: 'Todos los casos de *P* son casos de *Q*'; un enunciado de probalilidad estadística del tipo más simple tiene esta forma: 'La probabilidad de que un caso de *P* sea un caso de *Q* es *r*'. Mientras que el primero implica una afirmación acerca de cualquier caso particular de *P* —que es también un caso de *Q*—, el último no implica una afirmación similar acerca de cualquier caso particular de *P* o aun de cualquier conjunto finito de tales casos.[6] Esta circunstancia da lugar a la segunda característica distintiva mencionada antes: el enunciado *E* que describe el fenómeno sujeto a explicación, predicción o retrodicción (por ejemplo, que Juan contrajera el sarampión) no es lógicamente deducible de los enunciados explicativos aducidos [por ejemplo, (C_1) Juan estuvo expuesto al contagio del sarampión; (C_2) Juan no había tenido previamente el sarampión; (*L*) para las personas que no han tenido el sarampión y están expuestas al contagio, la probabilidad de que contraigan la enfermedad es de 0,92]. En cambio, suponiendo que los enunciados explicativos aducidos sean verdaderos, es muy probable, aunque no seguro, que *E* también sea verdadero. Este tipo de argumento, por lo tanto, es de carácter inductivo más bien que estrictamente deductivo: propone la conclusión *E* sobre la base de otros enunciados que, en caso de apoyarla fuertemente, sólo constituyen fundamentos parciales de esa conclusión. A un argumento de este tipo —ya sea que se lo use para explicación, predicción, retrodicción o aun con otros propósitos— se lo llamará *sistematización inductiva.* En particular, supondremos ante una sistematización inductiva que las premisas no implican lógicamente a la conclusión.[7] También se llamará sistematización inductiva al procedimiento para formular un argumento del tipo que se acaba de describir.

A modo de ilustración adicional, permítasenos registrar aquí dos argumentos explicativos que son del tipo inductivo que acabamos de caracterizar. Los propone Von Mises al afirmar que la noción corriente de explicación causal se ajustará eventualmente a cambios en la forma lógica de las teorías científicas (especialmente al uso de enunciados de probabilidad estadística como principios explicativos): "Pensamos", dice Von Mises, "que la gente irá aceptando en forma gradual enunciados causales de este tipo: *Porque* el dado está cargado el 'seis' aparece con más fre-

[6] Para una exposición más completa de este punto, véase, por ejemplo, Nagel (1939, sección 7), Reichenbach (1949, secciones 63-67), Cramér (1946, cap. 13).

[7] El uso explicativo y predictivo de las leyes estadísticas constituye quizás el tipo más importante de sistematización inductiva, pero nuestro concepto general de sistematización inductiva no requiere que aparezcan leyes de ese tipo entre las premisas. A decir verdad, como lo ha señalado Carnap (1950, págs. 574-575), a veces es posible hacer predicciones acerca de un conjunto finito de casos particulares, sin la mediación de ley alguna. Por ejemplo, las informaciones siguientes: 1) que se ha examinado una muestra amplia de casos de *P*, y 2) que todos sus elementos tienen la característica *Q*, y 3) que cierto caso *x no incluido en* la muestra es un caso de P apoyarán con fuerza la predicción de que *x* tiene también la característica *Q*. Asimismo, a veces es posible basar una sistematización inductiva sobre un conjunto de premisas que incluyen uno o más enunciados estrictamente universales pero que no incluyen leyes estadísticas. Se encontrará un ejemplo de esta sistematización en la sección 9, en la predicción basada sobre las fórmulas (9.6) — (9.12).

cuencia (pero no sabemos cuál será el resultado de la próxima jugada); o: *Porque* se elevó el vacío y se aumentó el voltaje, la radiación se tornó más intensa (pero no sabemos el número preciso de chispas que aparecerá en el próximo minuto)".[8] Resulta claro que ambos enunciados pueden considerarse como explicaciones inductivas de ciertos fenómenos físicos.

Todos los casos de sistematización científica que hemos considerado comparten esta característica: hacen uso de leyes generales o de principios generales, ya sea de forma estrictamente universal o de forma estadística. Estas leyes generales tienen la función de establecer conexiones sistemáticas entre hechos empíricos de tal modo que con su ayuda sea posible inferir a partir de algunos sucesos empíricos, otros sucesos semejantes a modo de explicación, predicción o retrodicción.

Cuando decimos en una explicación que el acontecimiento que describe E tuvo lugar "debido a" las circunstancias detalladas en C_1, $C_2 \ldots C_k$, esa oración es significativa si se puede referir a leyes generales que hacen que C_1, $C_2 \ldots C_k$ sean relevantes respecto de E en el sentido de que, una vez supuesta la verdad de las primeras, la verdad de la última resulta o bien cierta (como en una sistematización deductiva) o bien inductivamente probable (como en una sistematización inductiva). Por esta razón es de importancia crucial en las ciencias empíricas formular leyes generales.

2. *Observables y entidades teóricas*

La sistematización científica se propone esencialmente establecer un orden explicativo y predictivo entre los complejísimos "datos" de nuestra experiencia, o sea, entre los fenómenos que podemos "observar directamente". Cabe destacar, en consecuencia, que los mayores avances en la sistematización científica no se han llevado a cabo por medio de leyes referidas en forma explícita a *observables,* o sea, a cosas y acontecimientos que se pueden constatar por observación directa sino por medio de leyes que hacen referencia a diversas *entidades hipotéticas* o *teóricas*, o sea, objetos, acontecimientos y atributos que suponemos pero que no podemos percibir ni observar directamente de modo alguno.

Para una exposición más completa de este punto, será útil hacer referencia a la distinción, familiar aunque poco sutil, entre los dos niveles de la sistematización científica: el nivel de la *generalización empírica,* y el nivel de la *formación de teorías.*[9] Las etapas más tempranas en el desa-

[8] Mises (1951, pág. 188). Quizá no sea aconsejable considerar que son causales las explicaciones de este tipo: puesto que la concepción clásica de la causalidad está íntimamente ligada a la idea de leyes estrictamente universales que relacionan causa y efecto, sería mejor reservar el término 'explicación causal' para algunos argumentos explicativos que tienen la forma (1.1) en los cuales todas las leyes aducidas son de forma estrictamente universal.

[9] Northrop (1947, caps. III y IV), por ejemplo, presenta esta distinción de un modo muy interesante: se refiere a los dos niveles: como "la etapa de investigación como historia natural" y "la etapa de la teoría formulada en forma deductiva". Se encontrará una exposición lúcida y concisa sobre la idea en cuestión en Feigl (1948).

rrollo de una disciplina científica pertenecen generalmente al primer nivel, que se caracteriza por la búsqueda de leyes (de forma universal o estadística) que establecen conexiones entre los aspectos del tema en estudio que son directamente observables. Las etapas más avanzadas pertenecen al segundo nivel, en el que la investigación se dirige a la búsqueda de leyes comprensivas, en términos de entidades hipotéticas, que darán cuenta de las uniformidades establecidas en el primer nivel. En éste encontramos las generalizaciones corrientes de la física tales como 'Donde hay luz hay calor', 'El hierro se enmohece con el aire húmedo', 'En el agua la madera flota y el hierro se hunde'; pero también podríamos incluir leyes cuantitativas más precisas tales como las de Snell, Hooke, Kepler y Galileo, así como las generalizaciones de la botánica y de la zoología acerca de la concomitancia de ciertas características observables anatómicas, físicas, funcionales y otras en los miembros de una especie dada; generalizaciones de la psicología que afirman correlaciones entre diferentes aspectos observables del aprendizaje, de la percepción, etcétera, y diversas generalizaciones descriptas de la economía, la sociología y la antropología. Todas esas generalizaciones, ya sean de forma estrictamente universal o de forma estadística, pretenden expresar conexiones regulares entre fenómenos directamente observables, y se prestan por lo tanto, para un uso explicativo, predictivo y retrodictivo.

En el segundo nivel, encontramos enunciados generales que hacen referencia a campos eléctricos, magnéticos y gravitacionales, a moléculas, a átomos y a una variedad de partículas subatómicas; o al yo, al ello, al superyó, a la libido, a la sublimación, a la fijación y a la transferencia; o a las diversas entidades que invocan las recientes teorías del aprendizaje y que no son directamente observables.

De acuerdo con la distinción que hemos hecho aquí, supondremos que el vocabulario extralógico de la ciencia empírica o de cualquiera de sus ramas, se divide en dos clases: *términos observacionales* y *términos teóricos*. Con respecto a un término observacional es posible decidir, en circunstancias adecuadas, si el término se aplica o no a una situación dada por medio de observación.

La noción de observación puede interpretarse aquí en forma tan amplia que incluya no sólo la percepción sino también la sensación y la introspección, o se la puede circunscribir a la percepción de lo que en principio se puede constatar públicamente, o sea, de lo que también otros pueden percibir. La exposición que sigue será independiente de la mayor o menor amplitud con que se interprete la noción de observación. Vale la pena consignar, sin embargo, que la ciencia empírica apunta a un sistema de enunciados que se puedan probar públicamente; conforme con esto, se considera que los datos observacionales, cuya predicción correcta es la característica distintiva de una teoría exitosa, deben expresarse en términos tales que individuos diferentes puedan concordar en alto grado al decidir, por medio de la observación directa, si son aplicables en una situación dada o no lo son. Los enunciados que pretenden describir lecturas de instrumentos de medición, cambios en el color u olor que acompañan a una reacción química, verbalizaciones u otros tipos de compor-

tamiento manifiesto de un sujeto dado en determinadas condiciones observables, todos ellos ilustran el uso de los términos observacionales *intersubjetivamente aplicables.*[10]

Los términos teóricos, en cambio, habitualmente pretenden hacer referencia a entidades que no son directamente observables y a sus características; operan en las teorías científicas que intentan explicar las generalizaciones empíricas de una manera que pronto se examinará más de cerca.

Es obvio que la caracterización precedente de los dos vocabularios resulta vaga; no ofrece un criterio preciso por medio del cual se pueda clasificar inequívocamente cualquier término científico como observacional o teórico. Pero no se necesita aquí un criterio preciso de ese tipo; las cuestiones que se han de examinar en este ensayo son independientes del lugar exacto en que debe trazarse la línea divisoria entre los términos de ambos vocabularios: observacional y teórico.

3. *¿Por qué términos teóricos?*

El uso de términos teóricos en la ciencia da origen a un problema complicado. ¿Por qué debe recurrir la ciencia a la suposición de entidades hipotéticas cuando está interesada en establecer conexiones explicativas y predictivas entre observables? ¿No sería suficiente, y menos complicado para ese propósito, buscar un sistema de leyes generales que no mencionaran sino observables y estuvieran así expresadas sólo en términos del vocabulario observacional?

A decir verdad, se han formulado muchos enunciados en términos de observables; constituyen las generalizaciones empíricas mencionadas en la sección precedente. Pero, por desgracia, muchos de ellos, si no todos, tienen inconvenientes definidos; en general, su dominio de aplicación es muy limitado; y aun dentro de ese dominio, tienen excepciones, de modo tal que no son en realidad enunciados generales verdaderos. Consideremos, por ejemplo, uno de nuestros casos ilustrativos anteriores:

(3.1) En el agua la madera flota y el hierro se hunde.

[10] En su ensayo sobre el análisis del aprendizaje hecho por Skinner (en Estes y otros, 1945), Verplanck señala, en forma tangencial pero esclarecedora, cuán importante es el vocabulario observacional (los términos del lenguaje-de-los-datos,* como él los llama), que haya un alto grado de uniformidad en el uso de los términos entre diferentes experimentadores. Verplanck arguye que, a pesar de que gran parte del lenguaje-de-los-datos de Skinner es aceptable a este respecto, "lo contamina" la inclusión de dos tipos de términos que son inadecuados para la descripción de datos científicos objetivos. El primero incluye términos "que no pueden ser usados con éxito por muchos otros"; el segundo incluye ciertos términos a los que con propiedad se debería tratar como expresiones teóricas de un orden más elevado.

El carácter pragmático e impreciso del requisito de uniformidad intersubjetiva en el uso de los términos se pone de manifiesto en la conjetura de Verplanck "de que si uno tuviera que trabajar con Skinner, y leer sus informes con él, se sentiría capaz de hacer sus mismas distinciones y eventualmente de atribuir a algunos de sus términos la condición de lenguaje-de-los-datos" (*loc. cit.*, página 279n).

* En inglés, *data-language* [T].

Este enunciado tiene un dominio de aplicación estrecho en el sentido de que se refiere sólo a objetos de madera y de hierro y al hecho de flotar o de hundirse sólo en el agua.[11] Y lo que es aun más grave, tiene excepciones: ciertos tipos de madera se hundirán en el agua y una esfera hueca de hierro de dimensiones adecuadas flotará sobre ella.

Como lo muestra la historia de la ciencia, a menudo se pueden subsanar defectos de ese tipo atribuyendo a los fenómenos que se estudian ciertas características adicionales que, aunque no sean pasibles de observación directa, están relacionadas de un modo determinado con sus aspectos observables y que hacen posible establecer conexiones sistemáticas entre los últimos. Se obtiene una generalización mucho más satisfactoria que (3.1) por medio del concepto de gravedad específica de un cuerpo x, que es definible como el cociente de su peso y su volumen:

(3.2) Def. $e(x) = p(x) / v(x)$

Supongamos que se ha caracterizado a p y a v operacionalmente o sea, en términos de los resultados directamente observables de procedimientos de medición y que, por lo tanto, se los cuenta entre los observables. Entonces, podemos considerar a a, tal como se lo ha determinado por (3.2), como una característica que es observable menos directamente y la clasificaremos, sólo con el fin de que el ejemplo resulte simple, como una entidad hipotética. Podemos ahora establecer para e la siguiente generalización, que es un corolario del principio de Arquímedes:

(3.3) Un cuerpo sólido flota en un líquido si su gravedad específica es menor que la del líquido.

Este enunciado elude, en primer lugar, las excepciones que hemos consignado antes como refutatorias (3.1); predice correctamente el comportamiento de un pedazo de madera pesada y el de una esfera hueca de hierro. Por añadidura, tiene un alcance mucho más amplio: se refiere a cualquier tipo de objeto sólido y al hecho de flotar o de hundirse con respecto a un líquido cualquiera. Por supuesto, la nueva generalización tiene aún ciertas limitaciones e invita así a una nueva mejora. Pero, en lugar de continuar este proceso, examinemos ahora más de cerca el modo en que se lleva a cabo una conexión sistemática entre observables por medio de la ley (3.3), que implica incursiones en el dominio de los inobservables.

Supongamos que se desea predecir si cierto objeto sólido c flotará o se hundirá en un cuerpo dado l de líquido. Tendremos entonces que averiguar primero, por medio de procedimientos operacionales apropia-

[11] Sin embargo, se debería hacer notar que aquí se usa la idea de dominio de aplicación de una generalización en un sentido intuitivo que sería difícil de explicar. Se podría sostener, por ejemplo, que el dominio de aplicación de (3.1) es más estrecho de lo que se ha indicado: se podría interpretar que consta sólo de objetos-de-madera-colocados-en-el-agua y de objetos-de-hierro-colocados-en-el-agua. Por otra parte, puede proponerse como enunciado equivalente a (3.1) el que sigue: cualquier objeto tiene dos de estas propiedades: no ser de madera o bien flotar en el agua y no ser de hierro o bien hundirse en el agua. De esta forma, puede decirse que la generalización tiene el dominio de aplicación más amplio posible: la clase de todos los objetos.

dos, el peso y el volumen de c y de l. Los resultados de estas mediciones estarán expresados en los cuatro enunciados siguientes O_1, O_2, O_3, O_4:

(3.4)
$$O_1: \quad p(c) = p_1: \qquad O_2: \quad v(c) = v_1$$
$$O_3: \quad p(l) = p_2; \qquad O_4: \quad v(l) = v_2$$

donde p_1, p_2, v_1, v_2, son números reales positivos. Por medio de la definición 3.2), podemos inferir de (3.4), las gravedades específicas de c y de l.

(3.5) $e(c) = p_1/v_1$: $e(l) = p_2/v_2$

Supongamos ahora que el primero de esos valores es menor que el segundo; luego (3.4), por (3.5), implica que:

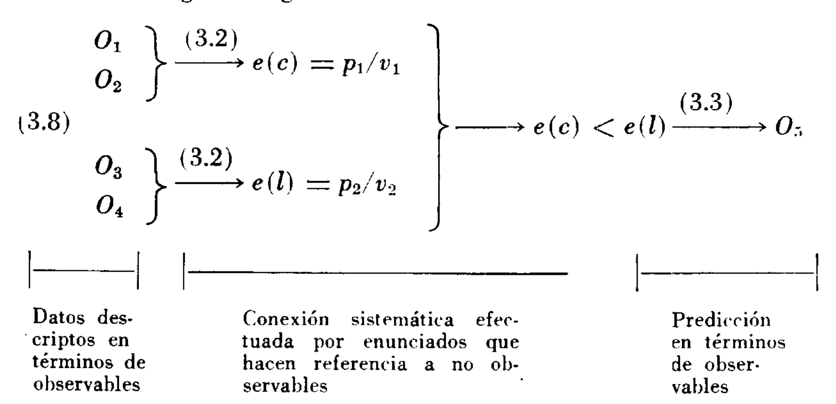
(3.6) $e(c) < e(l)$

Por medio de la ley (3.3) podemos ahora inferir que:

(3.7) c flota en l

A esta oración la llamaremos O_5. Luego O_1, O_2, O_3, O_4, O_5 comparten la característica de estar expresadas enteramente en términos del vocabulario observacional; porque en nuestra suposición 'p' y 'v' son términos observacionales, y también lo son 'c' y 'l' que nombran a ciertos cuerpos observables; finalmente 'flota en' es un término observacional porque en circunstancias adecuadas, la observación directa mostrará si un objeto observable dado flota en un líquido observable dado o si se hunde en él. Por otra parte, las oraciones (3.2), (3.3), (3.5) y (3.6) carecen de esa característica, porque todas contienen el término e que, en nuestro ejemplo, pertenece al vocabulario teórico.

La transición sistemática de los "datos observacionales" enumerados en (3.4) a la predicción (3.7) de un fenómeno observable está esquematizada en el diagrama siguiente:



Aquí, una flecha representa una inferencia deductiva; una oración encima de una flecha indica que se realiza la deducción por intermedio de ella, o sea, que la conclusión formulada a la derecha se sigue lógicamente de las premisas que figuran a la izquierda, tomadas en conjunción con la oración que está encima de la flecha. Notamos que ese argumento ilustra el esquema (1.1) donde O_1, O_2, O_3, O_4 constituyen los enunciados referidos a hechos pertinentes, las oraciones (3.2) y (3.3) ocupan el lugar de las leyes generales y O_5 el de E.[12]

Así, la suposición de entidades no observables sirve a los propósitos de la sistematización: proporciona conexiones entre observables en la forma de leyes que contienen términos teóricos y esta incursión en el dominio de entidades hipotéticas ofrece ciertas ventajas, algunas de las cuales ya se han indicado antes.

En el caso de nuestro ejemplo, sin embargo, una breve reflexión mostrará que las ventajas obtenidas con la "incursión teórica" podrían haberse obtenido igualmente sin emplear término teórico alguno. En virtud de la definición (3.2), la ley (3.3) puede reformularse como sigue:

(3.3) Un cuerpo sólido flota en un líquido si el cociente de su peso y su volumen es menor que el cociente correspondiente del líquido.

Es evidente que esta nueva versión participa de las ventajas que tiene (3.3) con respecto a la tosca generalización (3.1) y, por supuesto, permite la transición deductiva de O_1, O_2, O_3, O_4 a O_5, tal como lo hace (3.3) en conjunción con (3.2).

Se suscita, por lo tanto, esta cuestión: ¿Es posible en todos los casos encontrar para la sistematización llevada a cabo mediante principios generales que contienen términos teóricos, una formulación equivalente con enunciados generales expresados exclusivamente en términos observacionales? Como preparación para un examen de este importante problema, debemos considerar primero más de cerca la forma y la función de una teoría científica.

4. *Estructura e interpretación de una teoría*

Desde el punto de vista formal se puede considerar a una teoría científica como una conjunto de oraciones expresadas en términos de un vocabulario específico; se entenderá que el vocabulario V_T de una teoría

[12] Puesto que se presentó a (3.2) como una definición, podría considerarse inapropiado incluirla entre las leyes generales que realizan la transición predictiva de O_1, O_2, O_3, O_4 a O_5. A decir verdad, es posible concebir la deducción lógica aplicada en (1.1) de modo tal que incluya el uso de cualquier definición como premisa adicional. En este caso (3.3) es la única ley que se aduce en la predicción considerada aquí. Por otra parte, es posible también tratar a las oraciones tales como (3.2), que habitualmente se incluyen entre las que son sólo definiciones, en pie de igualdad con otros enunciados de forma universal, a los que se clasifica como leyes generales. Es favorable a este enfoque, por ejemplo, la consideración de que cuando una teoría entra en conflicto con los datos empíricos pertinentes, son algunas veces las "leyes" y otras las "definiciones" las que se modifican para ajustarse a los elementos de prueba. Nuestro análisis de la sistematización deductiva es neutral en lo que respecta a este problema.

T consta de los términos extralógicos de *T*, o sea de aquellos que no pertenecen al vocabulario de la lógica pura. Habitualmente, se definen algunos términos de V_T por medio de otros; pero so pena de incurrir en circularidad o en un regreso al infinito, no se puede definir así a todos ellos. En consecuencia, se puede suponer que *V* está dividido en dos subconjuntos: *términos primitivos* —aquellos para los que no se especifica una definición— y *términos definitivos*. Análogamente, muchas de las oraciones de una teoría son derivables de otras por medio de los principios de la lógica deductiva (y de las definiciones de los términos definidos); pero, so pena de incurrir en círculo vicioso o regreso al infinito en la deducción, no pueden fundamentarse así todas las oraciones de la teoría. En consecuencia, el conjunto de oraciones que componen *T* se divide en dos subconjuntos: *oraciones primitivas* o *postulados* (también llamados *axiomas*) y *oraciones derivadas* o *teoremas*. De aquí en adelante supondremos que las teorías se presentan en la forma de sistemas axiomatizados como los que aquí se describen: primero, enuncian los términos primitivos, los derivados y las definiciones de estos últimos; y segundo, enuncian los postulados. Además, consideraremos que están formuladas dentro de un marco lingüístico cuya estructura lógica claramente especificada determina, en particular, las reglas de la inferencia deductiva.

Los paradigmas clásicos de los sistemas deductivos de este tipo son las axiomatizaciones de varias teorías matemáticas, tales como la geometría euclídea y las diversas formas de geometrías no euclídeas, la teoría de los grupos y otras ramas del álgebra abstracta;[13] pero ahora se ha dado igualmente una forma axiomática o aproximaciones a ésta a cierto número de teorías de la ciencia empírica; entre esas teorías se encuentran partes de la mecánica clásica y relativista[14] y de la teoría biológica[15] y algunos sistemas teóricos en psicología, especialmente en el terreno del aprendizaje;[16] en la teoría económica, el concepto de utilidad entre otros, ha recibido tratamiento axiomático.[17]

Si se han especificado los términos primitivos y los postulados de un sistema axiomatizado pueden ponerse a prueba los teoremas, o sea, es

[13] Se podrá encontrar una lúcida exposición elemental de la naturaleza de los sistemas matemáticos axiomatizados en Cohen y Nagel (1934, cap. VI; también reimpreso en Feigl y Brodbeck (1953). Para un análisis similar, con énfasis especial sobre la geometría, véase también Hempel (1945). Una excelente descripción sistemática del método axiomático es la que proporciona Tarski (1941, cap. VI-X); esta presentación, que utiliza algunos conceptos de la lógica simbólica elemental, desarrollados en capítulos anteriores, incluye varios ejemplos sencillos de la matemática. Braithwaite, en los tres primeros capítulos de su obra publicada en 1953, lleva a cabo un estudio lógico cuidadoso de los sistemas deductivos en la ciencia empírica, prestando especial atención al papel desempeñado por los términos teóricos, y Woodger, especialmente en sus obras publicadas en 1937 y 1939, da una exposición más avanzada desde el punto de vista lógico del método axiomático, unido con aplicaciones a la teoría biológica.

[14] Véanse por ejemplo, Hermes (1938), Walker (1943-1949), McKinsey, Sugar y Suppes (1953), McKinsey y Suppes (1953, Rubin y Suppes (1953), y las referencias adicionales que proporcionan esas publicaciones. Un importante trabajo pionero en ese terreno es el de Reichenbach (1924).

[15] Véase en especial Woodger (1937) y (1939).

[16] Véase por ejemplo, Hull y otros (1940).

[17] Por ejemplo, en von Neumann y Morgenstern (1947), cap. III y apéndice.

posible derivar nuevos enunciados a partir de los primitivos por medio de los cánones puramente formales de la lógica deductiva, sin ninguna referencia a los significados de los términos y oraciones en cuestión; a decir verdad, para el desarrollo deductivo de un sistema axiomatizado no se necesita para nada asignar significado a sus expresiones, sean éstas primitivas o derivadas.

Sin embargo, un sistema deductivo sólo puede funcionar como teoría en la ciencia empírica si se le ha dado una *interpretación* con referencia a fenómenos empíricos. Puede llevarse a cabo esta interpretación por medio de la especificación de un conjunto de *oraciones interpretativas* que relacionan ciertos términos de vocabulario teórico con términos observacionales.[18] Se examinará en detalle el carácter de esas oraciones en las secciones siguientes; por ahora puede mencionarse como ejemplo que las oraciones interpretativas podrán tomar la forma de las definiciones llamadas operacionales, es decir, de enunciados que especifican el significado de términos teóricos con la ayuda de términos observacionales; son de especial importancia entre éstas las reglas para la medición de cantidades teóricas por medio de respuestas observables de instrumentos de medición o de otros indicadores.

Se puede esbozar ahora, mediante un ejemplo, la manera en que una teoría establece conexiones explicativas y predictivas entre enunciados expresados en términos observacionales. Supongamos que se use la teoría de la mecánica de Newton para estudiar los movimientos de los cuerpos bajo la influencia exclusiva de su atracción gravitacional mutua, siendo estos cuerpos los componentes de un sistema biestelar, o la luna y un cohete que costea libremente la superficie lunar cien millas por encima de ésta. Sobre la base de datos observacionales apropiados, se pueda asignar a cada uno de los dos cuerpos una cierta masa, y en un instante dado, t_0, una cierta posición y velocidad dentro de un marco determiado de referencia. Así, se da un primer paso que conduce, por intermedio de oraciones interpretativas en forma de reglas de medición, desde ciertos enunciados O_1, O_2, ... O_k que describen las lecturas obsrevables de los instrumentos, hasta ciertos enunciados teóricos, H_1, H_2, ... H_6, que asignan a cada uno de los cuerpos un valor numérico específico para las cantidades teóricas: masa, posición y velocidad. A partir de estos enunciados, la ley

[18] En las obras sobre metodología, los enunciados que dan una interpretación empírica de los términos teóricos han recibido una variedad de nombres. Por ejemplo, Reichenbach, que subrayó desde un principio la importancia de esta idea con referencia especial a la relación entre geometría pura y física, habla de *definiciones coordinativas* (1928, sección 4; también 1951, cap. VIII). Campbell [(1920, cap. VI; un extracto de este capítulo fue reimpreso en Feigl y Brodbeck (1953)] y Ramsey (1931, págs. 212-236) se refieren a un *diccionario* que conecta a los términos empíricos con los términos teóricos (véase también la sección 8). Margenau (1950, en especial, cap. 4) habla de *reglas de correspondencia*, y Carnap (1956) ha usado igualmente el término general '*reglas de correspondencia*'. Se puede considerar a las *correlaciones epistémicas* de Northrop (1947, en especial cap. VII) como un tipo especial de enunciados interpretativos. Para una exposición de la interpretación como procedimiento semántico, véase Carnap (1939, secciones 23, 24, 25) y Hutten (1956, en especial cap. II). Una exposición más completa de los enunciados interpretativos está incluida en las secciones 6, 7 y 8 de este capítulo.

de gravitación, que está expresada íntegramente en términos teóricos, conduce a otro enunciado teórico H_7 que especifica la fuerza de la atracción gravitacional que los cuerpos ejercen el uno sobre el otro en t_0; y H_7, en conjunción con los enunciados teóricos precedentes y las leyes de la mecánica de Newton implica, por intermedio de un argumento deductivo que comprende los principios del cálculo, ciertos enunciados H_8, H_9, H_{10}, H_{11}, que dan las posiciones y velocidades de los dos objetos en un momento posterior determinado, t_1. Finalmente el uso inverso de las oraciones interpretativas conduce desde los últimos cuatro enunciados teóricos hasta un conjunto de oraciones O'_1, O'_2, ... O'_m, que describen fenómenos observacionales, especialmente lecturas de instrumento que indican las posiciones y velocidades predichas.

Por medio de un esquema análogo a (3.8), puede representarse el proceso como sigue:

$$(4.1)\quad [O_1, O_2 \ldots O_k] \xrightarrow{R} [H_1, H_2 \ldots H_6] \xrightarrow{G} [H_1, H_2 \ldots H_6, H_7]$$

$$\xrightarrow{LM} [H_8, H_9, H_{10}, H_{11}] \xrightarrow{R} [O'_1, O'_2 \ldots O'_m]$$

Aquí, R es el conjunto de reglas de medición para la masa, posición y velocidad; esas reglas constituyen las oraciones interpretativas; G es la ley de gravitación de Newton, y LM las leyes del movimiento de Newton.

En lo que se refiere a la psicología, se han presentado repetidamente en las obras sobre temas metodológicos, análisis esquemáticos de la función de las teorías o de las hipótesis que comprenden "variables intervinientes" que son similares a los expuestos.[19] En éstos los datos observacionales con los que comienza el procedimiento conciernen por lo general a ciertos aspectos observables del estado inicial de un sujeto dado, más ciertos estímulos observables que actúan sobre él; y los enunciados observacionales finales describen una respuesta dada por el sujeto. Los enunciados teóricos que median en la transición de los primeros a los últimos se refieren a diversas entidades hipotéticas, tales como impulsos, reservas, inhibiciones u otras características, cualidades o estados psicológicos cualesquiera que postule la teoría en cuestión y no sean directamente observables.

5. *El dilema del teórico*

La exposición precedente sobre la función de las teorías suscita de nuevo el problema que encontramos en la sección 3, si no se puede evitar por entero la incursión teórica en un dominio de cosas, acontecimientos o características que no son directamente observables. Supongamos, por ejemplo, que —como suele ocurrir— tanto las oraciones interpretativas como las leyes enunciadas en la teoría tengan la forma de ecuaciones que relacionan ciertas expresiones en términos de cantidades teóricas ya sea

[19] Podrá encontrarse una lúcida y concisa exposición, por ejemplo, en Bergmann y Spence (1941) .

con otras expresiones semejantes o con expresiones en términos de cantidades observables. Se puede plantear entonces el problema de la manera sucinta en que lo formula Hull: "Si usted tiene un encadenamiento de ecuaciones confiables que se extiende desde las condiciones antecedentes que son observables hasta las condiciones consecuentes que también lo son, ¿por qué usar varias ecuaciones en el caso en que bastaría una, aun cuando usar varias pudiera no ser positivamente pernicioso?" [20] Skinner hace la misma observación en forma más general cuando critica en las teorías psicológicas la construcción de cadenas causales en las que un primer eslabón, que consiste en un acontecimiento observable y controlable se conecta con un eslabón final (tercero) del mismo tipo por medio de un eslabón intermediario que, por lo común, no está sujeto a observación y control. Skinner argumenta: "A menos que haya un punto débil en nuestra cadena causal de modo tal que el primer eslabón no determine legalmente al segundo, o el segundo al tercero, el primero y el tercero deben estar legalmente relacionados. Si debemos retroceder siempre más allá del segundo eslabón para la predicción y el control, podremos evitar muchas digresiones tediosas y agotadoras considerando al tercer eslabón como una función del primero".[21]

La conclusión que estos argumentos sugieren podría llamarse la *paradoja del teorizar.* Afirma que si los términos y principios generales de una teoría científica sirven a su propósito, es decir, si establecen conexiones definidas entre fenómenos observables, se puede prescindir de ellos puesto que cualquier cadena de leyes y enunciados interpretativos que estableciera tal conexión sería reemplazable por una ley que encadenara directamente antecedentes observacionales a consecuencias observacionales.

Si añadimos a esta tesis crucial dos enunciados más que son obviamente verdaderos, obtenemos las premisas de un argumento con la clásica forma de un dilema:

(5.1) Si los términos y principios de una teoría sirven a su propósito son innecesarios, como ya se lo ha señalado, y si no sirven a su propósito son sin duda innecesarios. Pero en una teoría cualquiera, o bien sus términos y principios sirven a su propósito o bien no lo hacen. En consecuencia, los términos y principios de una teoría cualquiera son innecesarios.

A este argumento, cuya conclusión concuerda con los puntos de vista de los conductistas metodológicos extremos en psicología se lo llamará *el dilema del teórico.*

Sin embargo, antes de dejarnos llevar por la alegría o la tristeza que nos puede proporcionar el resultado de este argumento, será conveniente recordar que las consideraciones que se adujeron hasta aquí para apoyar su primera premisa crucial fueron formuladas en forma muy esquemática. Para formarnos un juicio más fundamentado sobre el problema, será necesario, pues, averiguar si se puede completar el esquema de modo tal que dé lugar a un argumento convincente. A esta tarea nos dedicaremos ahora.

[20] Hull (1943, pág. 284).
[21] Skinner (1953, pág. 35).

6. *Definiciones operacionales y oraciones reductivas*

Será conveniente empezar por considerar más de cerca el carácter de las oraciones interpretativas. En el caso más simple, una oración de ese tipo podría ser una *definición explícita* de una expresión teórica en términos de expresiones observacionales, como lo ilustra (3.2). En este caso, el término teórico es innecesario en el sentido de que siempre se lo puede evitar reemplazándolo por una expresión observacional, su definiens. Si se definen así todos los primitivos de una teoría *T*, resulta claro que se puede formular *T* enteramente en términos observacionales, de modo tal que todos sus principios generales sean leyes que relacionen directamente observables entre sí.

Esto sería verdad, en particular, para cualquier teoría que pudiera satisfacer los cánones del operacionismo en el sentido restringido siguiente: se introduce cada uno de los términos mediante una definición explícita que establece una respuesta observable, siendo esta última necesaria y suficiente, en determinadas condiciones de prueba observables, para la aplicabilidad del término en cuestión. Supongamos, por ejemplo, que el término teórico es un predicado de primer grado o un término de propiedad 'Q'. Entonces una definición operacional del tipo mencionado adoptaría la forma

(6.1) Def. $Qx \equiv (Cx \supset Ex)$

o sea, un objeto x tiene (por definición) la propiedad Q si y sólo si, en las condiciones de prueba de tipo C, muestra un efecto o respuesta de tipo E. La definición de Tolman de la expectativa de alimento proporciona un ejemplo: "Cuando afirmamos que una rata espera alimento en *L*, lo que afirmamos es que *si*: 1) ha sido privada de alimento, 2) se la ha adiestrado para que siga el sendero *S*, 3) se la coloca ahora en el sendero *S*, 4) el sendero *S* está bloqueado y 5) hay otros senderos que se alejan del sendero *S* y uno de ellos conduce directamente al lugar *L*, entonces la rata seguirá el sendero que conduce directamente al lugar *L*".[22] Se puede obtener esta formulación reemplazando en (6.1), 'Qx' por 'la rata x espera alimento en el lugar *L*', 'Cx' por la conjunción de las condiciones 1), 2), 3), 4), 5) para la rata x, y 'Ex' por 'x sigue el sendero que conduce directamente al lugar *L*'.

Sin embargo, como lo ha mostrado Carnap en un argumento ahora clásico,[23] esta manera de definir los términos científicos, por natural que pueda parecer, tropieza con una seria dificultad. En la interpretación extensional canónica, una proposición condicional, tal como el definiens en (6.1) es falsa sólo si su antecedente es verdadero y su consecuente falso. Luego, para cualquier objeto que no satisfaga las condiciones de prueba *C*, de modo tal que el antecedente del definiens sea falso, el definiens

[22] Tolman, Ritchie y Kalish (1946, pág. 15). Véase también el análisis crítico detallado de la caracterización de la expectativa que ha hecho Tolman en MacCorquodale y Meehl (1945, págs. 179-181).
[23] Véase Carnap (1936-1937), sección 4.

como un todo es verdadero; por consiguiente, se asignará a ese objeto la propiedad *Q*. En nuestro ejemplo tendríamos que decir que una rata, no sometida a las condiciones (1)-(5) ya enunciadas, esperará alimento en *L*, cualquiera sea su comportamiento.

Una salida para esta dificultad puede ser la siguiente: cuando decimos que una determinada rata espera alimento en *L*, nos proponemos atribuir al animal un estado o disposición que, en las circunstancias (1)-(5) ya enunciadas, será la causa de que la rata siga el sendero que conduce directamente a *L*; en conscuencia, en una definición operacional apropiada, *E* debe ligarse a *C* nomológicamente, o sea en virtud de leyes generales que expresen conexiones causales. El conectivo extensional 'si... entonces' —que no requiere una necesidad de conexión lógica o nomológica— tendrá por lo tanto, que ser reemplazado en (6.1) por una conexión nomológica más estricta que quizá pudiera formularse: 'si... entonces, con necesidad causal...' Sin embargo, la idea de necesidad causal o nomológica tal como se ha empleado aquí no es suficientemente clara en la actualidad como para que este enfoque resulte fructífero.[24]

Carnap[25] ha propuesto otro modo de eliminar la dificultad con la que se tropieza en las definiciones de la forma (6.1); consiste en proporcionar, en lugar de una especificación completa, una especificación parcial del significado de '*Q*'. Esto se hace por medio de las llamadas oraciones reductivas; en el caso más simple, (6.1) podría reemplazarse por la siguiente *oración reductiva bilateral*:

(6.2) $Cx \supset (Qx \equiv Ex)$

Esta oración especifica que si un objeto está en condiciones de prueba de tipo *C*, tiene la propiedad *Q* si y sólo si muestra una respuesta de tipo *E*. Aquí, el uso de conectivos extensionales no tiene ya los aspectos indeseables que mostraba en (6.1). Si un objeto no está en condiciones de prueba C, la fórmula íntegra (6.2) es verdadera pero eso no implica que el objeto tenga la propiedad *Q* o bien que no la tenga. Por otra parte, mientras que (6.1) ofrece una definición explícita completa de '*Q*', (6.2) especifica el significado de '*Q*' sólo parcialmente, a saber, sólo para los objetos que satisfacen la condición *C*; para los que no lo hacen, se deja el significado de '*Q*' sin especificar. En nuestro ejemplo, (6.2) especificaría el significado de '*x* espera alimento en *L*' sólo para las ratas que satisfacen las condiciones (1)-(5); para ellas seguir el sendero que conduce a *L* sería una condición necesaria y suficiente de la expectativa de alimento. En lo que se refiere a las ratas que no satisfacen las condiciones de prueba (1)-(5), el significado de '*x* espera alimento en *L*' quedaría abierto; se lo podría especificar más por medio de oraciones reductivas adicionales.

De hecho, ésta es la interpretación que requiere la noción de expec-

[24] Sobre este punto y sobre el problema general de elucidar el concepto de ley natural, véase Braithwaite (1953), cap. IX, Burks (1951), Carnap (1956), sección 9, Goodman (1955), Hempel y Oppenheim (1948), parte III, Reichenbach (1954).

[25] En su teoría de las oraciones reductivas, desarrollada por Carnap (1936-1937). Sin embargo, se plantea la cuestión de saber si ciertas condiciones que Carnap impone a las oraciones reductivas no incluyen en forma implícita modalidades causales. Sobre este punto, véase Hempel (1963) sección 3.

tativa de alimento propuesta por Tolman. Ya que mientras el pasaje que acabamos de citar parece tener exactamente la forma (6.1), esta interpretación queda excluida por la oración que sigue inmediatamente a la citada: "Cuando aseveramos que la rata no espera alimento en el lugar *L*, lo que decimos es que, en las mismas condiciones, no seguirá el camino que conduce al lugar *L*". La interpretación total que se da así a 'la rata *x* espera comida en *L*' se formula más satisfactoriamente en términos de una oración de la forma (6.2) siguiendo el esbozo propuesto en el párrafo precedente.[26]

Las oraciones reductivas ofrecen una formulación precisa del contenido de las definiciones operacionales, como lo ilustra claramente nuestro ejemplo. Al interpretar dichas definiciones como especificaciones de significado sólo parciales, se trata a los conceptos teóricos como si fueran "abiertos"; y al suministrar un conjunto de oraciones reductivas diferentes y mutuamente suplementarias para un término dado, se pone de manifiesto que, para muchos términos teóricos, se dispone de diferentes criterios operacionales de aplicación pertenecientes a contextos diferentes.[27]

Sin embargo, si bien un análisis en términos de oraciones reductivas considera a los términos teóricos como definidos en forma incompleta con referencia a observables, esto no prueba que *no se pueda* lograr una definición explícita completa de expresiones teóricas en términos observacionales. Y en verdad, parece cuestionable que pedir una *prueba* a ese efecto sea tan siquiera significativo. En la próxima sección se tratará ese problema con algún detalle.

7. *Definibilidad de los términos teóricos por medio de un vocabulario observacional*

La primera observación, bastante general, que haremos aquí es la siguiente: la definición de un término cualquiera, por ejemplo, 'v', por medio de un conjunto V de otros términos, por ejemplo, 'v_1', 'v_2' ... 'v_n' tendrá que especificar una condición necesaria y suficiente para la aplicabilidad de 'v' expresada en términos de algunos o de todos los miembros de V. Y a fin de que seamos capaces de juzgar si se puede proceder así en un caso determinado, tendremos que saber cómo se deben entender los términos en consideración. Por ejemplo, el vocabulario que consta de los términos 'varón' y 'vástago de' permite la formulación de una condición necesaria y suficiente de aplicación para el término 'hijo de' en su sentido biológico pero no en su sentido legal. Se puede indicar cómo se deben entender los términos dados, especificando un conjunto U de oraciones a las que se debe considerar como verdaderas, y que conectan cada uno de los términos dados con los demás y quizá con otros términos. Así, U será

[26] De hecho, los autores resumen la especificación total de significado que se realiza en los pasajes citados en su "definición" DF II, que tiene exactamente la forma (6.2) de una oración bilateral reductiva para 'la rata *x* espera comida en *L*'. [Tolman, Ritchie y Kalish (1946, pág. 15).]

[27] Para una exposición más completa, véase Carnap (1936-1937), sección 7 y Carnap (1956), sección 10.

un conjunto de oraciones que contendrá a 'v_1', 'v_2', ... 'v_n' y posiblemente también a otras constantes extralógicas. Por ejemplo, en el caso del uso biológico de los términos 'hijo', 'varón', 'vástago', con referencia a los seres humanos, puede darse el conjunto siguiente —llamémosle U_1— de oraciones: 'Todo hijo es varón', 'Ninguna hija es varón', '*x* es vástago de *y* si y sólo si *x* es hijo o hija de *y*'.

Por lo general las oraciones de *U* especifican qué suposiciones acerca de los conceptos en consideración se deben hacer cuando se busca una definición; y el problema de la definibilidad se transforma en el de la posibilidad de formular, en términos de v_1, v_2 ... v_n, una condición que, *en virtud de las suposiciones incluidas en U*, sea a la vez necesaria y suficiente para *v*. Así, aplicando una idea propuesta y desarrollada técnicamente por Tarski,[28] vemos que el concepto de definibilidad de '*v*' por medio de 'v_1', 'v_2', ... 'v_n', adquiere un significado preciso sólo si está relativizado explícitamente con referencia a un conjunto *U* de suposiciones especificadas. Ese significado preciso puede enunciarse ahora del siguiente modo:

(7.1) '*v*' es definible por medio del vocabulario $V = [\text{'}v_1\text{'}, \text{'}v_2\text{'}, \ldots \text{'}v_n\text{'}]$

relativo a un conjunto *U* finito de enunciados que contienen por lo menos a '*v*' y a todos los elementos de *V*, si de *U* es deducible por lo menos una oración que enuncie una condición necesaria y suficiente para *v*, en términos de constantes extralógicas que sean indefectiblemente miembros de *V*.

Si todos los términos en estudio son predicados uniposicionales de primer orden, por ejemplo, entonces una oración del tipo que se requiere podrá enunciarse de modo más simple:

(7.2) $v(x) \equiv D\,(x, v_1, v_2, \ldots v_n)$

donde la expresión del segundo miembro representa una función oracional cuya única variable libre es '*x*' y que no contiene otras constantes extralógicas que las incluidas en *V*.

De manera similar, en nuestro ejemplo, el conjunto *U*, especificado anteriormente implica el enunciado:

x es hijo de *y* $\equiv$ (*x* es varón y *x* es vástago de *y*)

de modo tal que, en relación con U_1, 'hijo' es definible como 'vástago masculino'.

Se considera, por lo general, que una definición, cuando no es simplemente una convención que introduce una oración abreviatoria (tal como la convención por la cual 'x^5' sustituye a '$x \cdot x \cdot x \cdot x \cdot x$'), establece la *sinonimia* de dos expresiones, o como se dice a menudo, la *identidad de sus significados*. Ahora bien, la cuestión de la definibilidad de un término dado, '*v*' por medio de un conjunto *V* de otros términos, sin duda no es sólo de orden notacional; se interpretará normalmente que tiene la posibilidad de expresar el significado del término '*v*' con relación a los significados

[28] Véase Tarski (1935, en especial págs. 80-83).

de los miembros de *V*. Si se concibe así a esta cuestión, la información necesaria para responder a ella concernirá naturalmente a los *significados* de '*v*' y de los miembros de *V*; de acuerdo con esto, no se requerirá sólo que los enunciados de *U* que proporcionan esta información sean verdaderos, sino que sean además analíticos, o sea, verdaderos en virtud de los significados atribuidos a los términos constitutivos. En este caso, los enunciados de *U* tendrán el carácter de postulados significativos en el sentido de Kemeny y Carnap.[29]

Sin embargo, en un estudio de la definibilidad de las expresiones teóricas por medio de términos observacionales, no es ni necesario ni aconsejable concebir la definición de esta manera intencional. Porque en primer lugar, la idea de significado, y las nociones afines a ella, tales como las de analiticidad y sinonimia no son tan claras como se pensó por largo tiempo[30] y será mejor, por lo tanto, eludirlas cuando sea posible hacerlo.

En segundo lugar, aunque se acepten esos conceptos como claramente inteligibles, no se puede concebir que la definibilidad de un término teórico resida exclusivamente en la existencia de una expresión sinónima que contenga sólo términos observacionales; sería suficiente si dispusiéramos de una expresión coextensa (más bien que una estrictamente cointensa, o sinónima) en términos de observables, ya que una expresión de este tipo enunciaría una condición observacional, empíricamente necesaria y suficiente, de aplicabilidad para el término teórico; y esto es todo lo que se requiere para nuestros propósitos. De hecho, se puede dar a la oración que enuncia la condición —que puede tener la forma (7.2), por ejemplo— el status de una verdad por definición, mediante una reformalización adecuada de la teoría.

Es interesante advertir aquí que se podría descubrir inductivamente una condición observacional necesaria y suficiente para un término teórico, por ejemplo '*Q*', aun cuando sólo dispusiéramos de una especificación parcial del significado de '*Q*' en términos de observables. Supongamos que, por ejemplo, se ha especificado un conjunto de condiciones diferentes de aplicación para '*Q*' por medio de oraciones reductivas bilaterales:

$$
\begin{array}{ll}
(7.3) & C_1x \supset (Qx \equiv E_1x) \\
 & C_2x \supset (Qx \equiv E_2x) \\
 & \dots\dots\dots\dots\dots \\
 & C_nx \supset (Qx \equiv E_nx)
\end{array}
$$

donde todos los predicados excepto '*Q*' son observacionales. Supóngase además que una investigación adecuada conduce a las siguientes generalizaciones empíricas:

[29] Véase Kemeny (1951) y (1952), Carnap (1952).

[30] Sobre este punto, véase en especial Quine (1951), Goodman (1949), White (1950) y (1956) parte II. Hay un examen crítico de la significación de la analiticidad con una referencia particular a los enunciados teóricos, por ejemplo en Pap (1953) y (1955) y en Hempel (1963). En los siguientes artículos, entre otros, se hallará la defensa de los conceptos de analiticidad y sinonimia: Carnap (1952) (1955), Grice y Strawson (1956), Martin (1952), Mates (1951), Wang (1955).

(7.4) $C_1x \supset (Ox \equiv E_1x)$
$C_2x \supset (Ox \equiv E_2x)$
. . . .
$C_nx \supset (Ox \equiv E_nx)$

donde 'O_x' representa una función oracional en 'x' que no contiene ningún término extralógico no-observacional. Estos descubrimientos en combinación con (7.3), aproyarían inductivamente la hipótesis:

(7.5) $Q_x \equiv O_x$

que presenta una condición observacional necesaria y suficiente para Q. Sin embargo, aunque (7.5) sea verdadera (su aceptación implica el 'riesgo inductivo" habitual), resulta claro que no expresa una sinonimia; si lo hiciera, no se necesitaría en primer lugar una investigación empírica para establecerla. Antes bien, afirma que, como cuestión de hecho, 'O' es coextenso con 'Q', o que O es una condición empíricamente necesaria y suficiente para Q. Y si lo deseamos, podemos imaginar que la teoría en cuestión más su interpretación está organizada en la forma de un sistema deductivo en el cual (7.5) se convierte en una verdad deficional, y (7.3) asume el carácter de un conjunto de enunciados empíricos equivalentes a los registrados en (7.4).

Se podría mencionar aquí, al pasar, que se requiere una interpretación extensional de la definibilidad igualmente amplia respecto de la posibilidad de "reducir" una disciplina científica, tal como la psicología a otra, tal como la biología, o aun la física y la química.[31] En efecto uno de los componentes de este problema es la posibilidad de definir los términos de la primera disciplina por medio de los de la segunda, y lo que se necesita para este propósito es nuevamente un conjunto de hipótesis empíricas que proporcionen, para cada término psicológico, una condición de aplicación necesaria y suficiente expresada en el vocabulario de la física y la química.

Cuando decimos, por ejemplo, que los conceptos de los diversos elementos químicos pueden definirse en términos físicos mediante una caracterización de los modos específicos en que sus moléculas se componen de partículas físicas elementales, es evidente que nos estamos refiriendo a los resultados de una investigación experimental más que a un mero análisis de lo que *significan* los términos que nombran a los diversos elementos. Si éste fuera el caso, resultaría poco comprensible que los problemas relacionados con la definibilidad de los términos científicos presentaran alguna dificultad y que fueran objeto de muchas conjeturas y controversias.

Las consideraciones precedentes tienen importantes implicaciones para la cuestión que nos ocupa, que es la de la posibilidad de definir todos los términos teóricos en la ciencia empírica mediante observables. Antes que nada, muestran que la cuestión, tal como se ha planteado, es elíptica: para completarla, tenemos que especificar un conjunto U de enunciados como

[31] Sobre el problema de "reducir" los conceptos de una disciplina a otra ofrecen importantes orientaciones: Nagel (1949) y (1951), Woodger (1952, pág. 271 y siguientes), Kemeny y Oppenheim (1956).

aquel al que se ha hecho referencia en (7.1). ¿Qué conjunto sería razonable elegir para ese propósito? Una elección natural recaería sobre el conjunto de todos los enunciados, en términos observacionales o teóricos, que la ciencia contemporánea acepta como probablemente verdaderos. Ahora bien, esta caracterización pragmático-histórica no es de ningún modo precisa y sin ambigüedades; hay una extensa área marginal que contiene enunciados para los cuales no se puede determinar con claridad si la ciencia contemporánea los acepta o si no lo hace. Pero cualquiera que sea la ubicación que se decida para los enunciados de esta área marginal, y cualquiera que sea el lugar en que —dentro de lo razonable— se trace el límite entre términos observacionales y teóricos, hay por lo menos una cuestión discutible: ¿el conjunto de los enunciados científicos actualmente aceptados implica una condición de aplicabilidad necesaria y suficiente para cada término teórico formulada mediante observables? Por cierto, los que han afirmado tal definibilidad no han apoyado su pretensión deduciendo realmente tales condiciones o presentando razones generales convincentes en favor de la posibilidad de hacerlo.

También se puede concebir la demanda de definibilidad como la afirmación de que será posible deducir eventualmente condiciones necesarias y suficientes del tipo requerido a partir de nuestro conocimiento científico a medida que éste se vuelve más amplio. (En este sentido se entiende habitualmente la definibilidad de los conceptos de la psicología en función de los de la biología o de la física y la química; pues parece claro que no pueden deducirse todos los enunciados de definiciones necesarios —aun en un sentido empírico, extensional— a partir de los principios psicológicos, biológicos, físicos y químicos corrientes.[32]) Pero afirmar la definibilidad de un término teórico en este sentido implica sostener: primero, que el término en cuestión no será abandonado en el desarrollo ulterior de la teorización científica; y segundo, que se descubrirán leyes generales que habrán de establecer ciertas condiciones necesarias y suficientes expresables en términos observacionales para la aplicación del término teórico en cuestión. Es evidente que no se puede fundamentar la verdad de ambas demandas por medio de argumentos filosóficos sino, en el mejor de los casos, mediante los resultados de la investigación científica que se lleve a cabo.

A pesar de lo precario del problema, filósofos de la ciencia y científicos interesados en metodología han presentado argumentos en pro y en contra de la posibilidad de definir términos teóricos con referencia a observables.

Algunos filósofos han insistido simplemente en que sólo lo que se acerca a una definición explícita en función de un vocabulario que se comprende con claridad puede proporcionar un método aceptable para introducir nuevos términos en el lenguaje de la ciencia; y argumentan que procediendo de otro modo los nuevos términos resultan ininteligibles.[33]

[32] Este punto está expuesto de manera más completa en Hempel (1951).

[33] Goodman es uno de los autores a quienes su "conciencia filosófica" obliga a aceptar esta posición (véase 1951, cap. II, sección I). Russell tomó una posición similar al insistir en que se debía concebir a los objetos físicos, como "construcciones

Volveremos sobre esta cuestión más adelante. Los representantes de esta posición no hacen una afirmación acerca de la definibilidad real de los términos teóricos que se usan en la ciencia empírica contemporánea; antes bien, acentúan la importancia de esclarecer las ideas de la ciencia reformulándolas, hasta donde sea posible, en un lenguaje con una estructura lógica, clara y simple, y de modo tal que todos los términos teóricos se introduzcan por medio de definiciones adecuadas.

Sin embargo, otros han sostenido que las teorías científicas y el modo en que funcionan tienen ciertas características lógicas o metodológicas generales a las que presumiblemente no afectan los cambios en el conocimiento científico, y que proporcionan una base para la cuestión de la definibilidad de los términos teóricos sin necesidad de examinar todos los enunciados que acepta la ciencia contemporánea o de esperar los resultados de la investigación que se lleve a cabo.

El argumento de Carnap, al que se ha hecho referencia en la sección 6, proporciona un ejemplo de ese tipo de procedimiento: muestra que las definiciones de la forma (6.1) no pueden servir para introducir los conceptos científicos que se debían especificar. Sin embargo, el argumento es limitado, en el sentido de que no muestra (y no pretende mostrar) que una definición explícita de los términos teóricos mediante observacionales sea, por lo general, imposible.

Recientemente[34] Carnap ha extendido su examen del problema en la siguiente dirección. Supongamos que un objeto dado *b* exhibe este tipo de comportamiento legal: toda vez que *b* está en condiciones de un tipo determinado observable, *C*, su respuesta es de un tipo determinado observable *E*. Decimos entonces que *b* tiene la disposición para reaccionar ante *C* por medio de *E*; para abreviar llamemos *Q* a esa propiedad disposicional. Es evidente que nuestra exposición anterior en la sección 6 concierne al problema de definir con precisión a '*Q*' en términos de '*C*' y '*E*'; habíamos advertido allí, siguiendo a Carnap, que tendríamos o bien que resignarnos a una especificación parcial del significado de '*Q*' por medio de la oración reductiva bilateral (6.2); o bien, si insistimos en una definición explícita completa, tendremos que usar modalidades nomológicas en el definiens.

Pero, cualquiera sea la que se elija entre estas vías alternativas, el término disposicional '*Q*' resultante tendrá la característica siguiente: si un objeto dado *b* está en una condición *C* y no da una respuesta *E*, o para abreviar, si Cb pero $\sim Eb$, esto establece de modo concluyente que *b* carece de la propiedad *Q*, o para abreviar, que $\sim Qb$. Carnap sostiene que esa característica distingue a los "términos disposicionales puros", tales como '*Q*', de los términos teóricos usados en la ciencia; porque a pesar de que estos últimos están relacionados con el vocabulario observacional por medio de ciertas oraciones interpretativas —a las que Carnap llama reglas de *C*—, esas reglas no permiten, en general, que un conjunto de datos observacionales (tales como 'Cb' y '$\sim Eb$') constituya un elemen-

lógicas" a partir de los datos de los sentidos, definibles así en términos de estos últimos. (Véase por ejemplo, 1929, cap. VIII).

[34] Véase Carnap (1956), en especial secciones 9, 10.

to de prueba concluyente en pro o en contra de la aplicabilidad del término teórico en una situación determinada. Hay dos razones para esta afirmación. Primero, las oraciones interpretativas para un término teórico dado proporcionan una interpretación observacional sólo dentro de un término limitado; así, por ejemplo, en el caso del término teórico 'masa' no hay ninguna regla de *C* directamente aplicable a una oración E_m tal que adscriba un cierto valor de masa a un cuerpo determinado si el valor es o bien tan pequeño que el cuerpo no es directamente observable o bien tan grande que el observador no puede manipular el cuerpo".[35]

Segundo, una interpretación observacional directa de un término teórico siempre implica el tácito reconocimiento de que la presencia o ausencia de la respuesta observable necesaria en una cierta situación de prueba debe servir como criterio sólo si no hay factores perturbadores o a condición de que "el estado del medio sea normal".[36] Así, por ejemplo, una regla de correspondencia podría determinar que la desviación de una aguja magnética es un síntoma observable de una corriente eléctrica en un alambre cercano, pero con el reconocimiento tácito de que la respuesta de la aguja debe contar sólo si no hay factores perturbadores, tales como, por ejemplo, una tormenta magnética repentina.

En términos generales, Carnap sostiene que "si un científico ha decidido usar cierto término '*M*' de modo tal que, para ciertas oraciones acerca de M, cualquier resultado observacional posible no pueda ser nunca un elemento de prueba concluyente sino, en el mejor de los casos, un elemento de prueba que permita un alto grado de probabilidad"; entonces el lugar apropiado para '*M*' es el vocabulario teórico.[37]

Deberíamos advertir, antes que nada, que si los argumentos de Carnap no están descaminados, prueban que no pueden concebirse los términos teóricos como disposicionales puros y así, aun cuando mediante el uso de modalidades nomológicas se lograran definiciones explícitas de estos últimos, el método no sería adaptable a los términos teóricos. Pero los argu-

[35] Carnap (1956), sección 10.

[36] Carnap (1950), sección 10.

[37] Carnap (1956), sección 10. Pap en sus obras de (1953) y (1955), secciones 10, 13 y 70 ha propuesto una idea semejante en espíritu a la de Carnap pero no tan clara en su contenido, con la pretensión (que Carnap no tuvo) de establecer la "imposibilidad de sostener la tesis de la definibilidad explícita" de los términos teóricos por medio de observacionales (Pap, 1953, pág. 8). Por otra parte, Bergmann afirma que muchos conceptos de la física teórica, incluyendo "aun las nociones de partícula de la física clásica, podrían introducirse mediante definiciones explícitas De paso, esto es verdad también en lo que respecta a los conceptos de la psicología científica", Bergmann (1951a, sección I). (En el mismo contexto, Bergmann advierte que el método de interpretación parcial parece necesario para disolver algunos de los enigmas de la teoría cuántica.) Sin embargo, esta fuerte aseveración se apoya principalmente sobre esbozos de algunas muestras de definiciones. Bergmann sugiere, por ejemplo, que se puede definir 'este lugar está en un campo eléctrico' mediante una oración de la forma 'si R_1, entonces R_2', donde R_1 *representa a la oración según* la cual hay un electroscopio en el lugar en cuestión, y R_2 representa a "la descripción del comportamiento de un electroscopio (en un campo eléctrico)" (1951, págs. 98-99). Sin embargo, puede cuestionarse a este tipo de definición sobre la base de los argumentos de Carnap, que acabamos de examinar. Además, algunos ejemplos, aunque son inobjetables, no pueden fundamentar la tesis general en discusión. Así, el problema permanece sin solución.

mentos no muestran —ni pretenden mostrar— que no se pueda definir explícitamente de ningún modo los términos teóricos en función de observables. De hecho, si se acepta la afirmación de Carnap citada en el párrafo precedente, muchos términos que se pueden definir explícitamente por medio del vocabulario observacional deberán ser calificados como teóricos. Por ejemplo, sea '*R*' un predicado observacional biposicional, y defínase del siguiente modo un predicado uniposicional 'M_1':

(7.6) Def. $M_1x \equiv (\exists y)\ Rxy$

o sea, un objeto x tiene la propiedad M_1 sólo en el caso en que está en relación R con por lo menos un objeto y. Si, por ejemplo, 'Rxy' representa 'x es menos pesado que y', M es la propiedad de ser inferior en el peso a por lo menos un objeto, o de no ser el más pesado de todos los objetos.

Supongamos, como es habitual hacerlo, que el dominio de los objetos en un estudio es infinito, o por lo menos, que no se le ha asignado un determinado número máximo de elementos. Consideremos ahora la posibilidad de elementos de prueba observacionales concluyen en pro o en contra de la oración 'M_1a' que atribuye M_1 a cierto objeto a. Es obvio que un solo descubrimiento observacional de que a tiene una relación R con cierto objeto b, o que Rab, bastaría para verificar 'M_1a' completamente. Pero ningún conjunto finito de datos observacionales: '$\sim Raa$' '$\sim Rab$', '$\sim Rac$', etc., bastaría para refutarla en forma concluyente. De acuerdo con el criterio de Carnap, por lo tanto, 'M_1', aunque definido en términos del predicado observacional '*R*', tendría quizá que ser clasificado como término teórico.

Pero es posible que en el pasaje antes citado Carnap se haya propuesto exigir de un término teórico '*M*' que para ciertas oraciones acerca de M ningún resultado observacional pueda constituir un elemento de prueba verificatorio o refutatorio concluyente. No obstante, aun los términos que satisfacen esa exigencia pueden definirse explícitamente en términos de observables. Sea '*S*' un predicado observacional triposicional; por ejemplo, '$Sxyz$' podría representar 'x está más lejos de y que de z'. Y defínase a 'M_2' como sigue:

(7.7) Def. $M_2x \equiv (\exists y)(z)[\sim (z = y) \supset Sxyz]$

En nuestro ejemplo, un objeto x tiene M_2 sólo cuando hay un objeto y respecto del cual x está más lejos que de cualquier otro objeto z. Consideremos ahora la oración 'M_2a'. Como se ve de inmediato, ningún conjunto finito de descubrimientos observacionales (todos los relevantes tendrían la forma '$Sabc$' o '$\sim Sabc$') puede constituir un elemento de prueba concluyente, ya sea verificatorio o refutatorio en relación a 'M_2a'. En consecuencia, a pesar de que se lo ha definido explícitamente en términos del predicado observacional '*S*', el término 'M_2' es teórico de acuerdo con el criterio sugerido por Carnap.

La exposición precedente ilustra un punto elemental pero importante: cuando se define un término, por ejemplo, un predicado uniposicional '*Q*' en términos de observables, su definiens debe enunciar una condición necesaria y suficiente para la aplicabilidad de '*Q*', o sea, para la verdad

de la forma '*Qb*'. Pero pese a que esa condición se enuncie por completo en términos observacionales, quizá no nos capacite todavía para decidir, sobre la base de un número finito de descubrimientos observacionales, si 'Q' se aplica a un objeto dado *b*, porque la condición de verdad de '*Qb*' tal como la proporciona el definiens puede no ser equivalente a una oración compuesta que sea función de verdad, cada uno de cuyos componentes expresa un descubrimiento observacional en potencia.

Para añadir un ejemplo más a los que se han propuesto antes: supongamos que el término de propiedad 'objeto de hierro' y los términos relacionales 'atrae a' y 'cercano a' están incluidos en el vocabulario observacional. En ese caso la definición:

(7.8) Def. x es un imán $\equiv x$ atrae a todo objeto cercano a él

consta de términos observacionales; pero el criterio que proporciona para saber si un objeto *b* es un imán no puede expresarse mediante un número finito cualquiera de descubrimientos observacionales; porque para establecer que *b* es un imán, tendríamos que mostrar que *b* atraerá a cualquier trozo de hierro que le acerquemos en un momento cualquiera; y esta afirmación se refiere a una infinidad de casos.

Para expresar la idea más formalmente, supongamos que nuestro vocabulario observacional contiene, además de nombre individuales para objetos observables, sólo predicados de primer orden de cualquier grado que representan atributos (o sea, propiedades o relaciones) observables en el sentido de que un pequeño número de observaciones directas bastará en condiciones adecuadas, para averiguar si un determinado objeto o grupo de objetos exhibe el atributo en cuestión.

Adoptemos ahora las siguientes definiciones: una *oración atómica* es una oración tal como '*Pa*', '*Rcd*', '*Sadg*', que adscribe un atributo observable a un determinado objeto o grupo de objetos. Una *oración básica* es una oración atómica o la negación de una oración atómica. Una *oración molecular* es una oración formada a partir de un número finito de oraciones atómicas por medio de conectivos oracionales. Se incluirán las oraciones básicas entre las moleculares.

Las oraciones básicas pueden considerarse como los enunciados más simples que describen resultados potenciales de la observación directa: afirman que un conjunto determinado de (uno o más) objetos posee tal o cual atributo observable o que carece de él.

Ahora bien, para todo enunciado molecular *E*, hay ciertas clases finitas de enunciados básicos que implican *E*, y otras clases que implican la negación de *E*. Así, ['*Pa*'] y también ['$\sim$ *Pa*', '*Rab*'], por ejemplo, implican la oración molecular '*Pa* v ($\sim$ *Pa* · *Rab*)'; mientras que el conjunto ['$\sim$ *Pa*', '$\sim$ *Rab*'] implica su negación. Esto muestra que para cada oración molecular *E*, es posible determinar un conjunto de oraciones básicas cuya verdad verificaría la negación de *E*, y así refutaría *E* de modo concluyente. Luego una oración molecular es pasible, "en principio" tanto de verificación observacional concluyente como de refutación observacional concluyente, en el sentido de que se pueden describir datos potenciales que, al actualizarse, verificarían la oración y otros que la refutarían; pero

no, por supuesto, en el sentido de que pudieran darse datos de ambos tipos a la vez, ya que son incompatibles entre sí.

Hay inclusive algunas oraciones de forma no molecular, por ejemplo, oraciones que contienen cuantificadores en forma no vacía, que son a la vez completamente verificables y completamente refutables, en el sentido que se acaba de aclarar.[38] Por ejemplo, la oración [‘*Qa*’] implica ‘$(x)\ (Px \vee Qa)$’ y las oraciones [‘$\sim Pb$’, ‘$\sim Qa$’] implican la negación de aquélla. Un argumento similar se aplica a ‘$(\exists x)\ (Px \cdot Qc)$’.

Sin embargo, por lo general las oraciones no moleculares no son a la vez verificables y refutables. Esto vale en particular para todas las oraciones no moleculares de forma puramente universal, o sea las que no contienen ninguna constante individual como ‘$(x)\ (Px \supset Qx)$’, pero es verdadero también para muchas oraciones cuantificadas que contienen constantes individuales. Así, si ‘*R*’ y ‘*S*’ son predicados observacionales, entonces oraciones del tipo ‘$(\exists y) Ray$’ no son refutables y oraciones del tipo ‘$(y)\ (\exists z) Sayz$’ y $(\exists y)\ (z) Sayz$’ no son ni verificables ni refutables, como ya se ha visto.

Las definiciones explícitas de los términos científicos mediante un vocabulario observacional pueden dividirse, de acuerdo con lo expuesto, en dos tipos: las que proporcionan *criterios de aplicación observacionales finitos* para el término definido y las que no lo hacen. Las primeras son simplemente aquellas cuyos definiens aplicados a un caso particular dan lugar a una proporción que es a la vez verificable y refutable. La siguiente definición:

(7.9) Def. Hijo $xy \equiv$ varón x · vástago xy

pertenece a ese tipo, ya que si se aplica el definiens a dos individuos particulares, por ejemplo *a* y *b*, da lugar a la oración:
‘varón *a* · vástago *ab*’, que es a la vez verificable y refutable y que proporciona así un criterio de aplicación observacional finito para aplicar el término ‘hijo’ a *a* en relación con *b*. Por otra parte, las definiciones anteriores (7.6), (7.7) y (7.8) están entre las que no ofrecen criterios de aplicación observacionales finitos para el término que definen, lo que ya se habían señalado antes.

Sin embargo, la circunstancia de que se introduzca originariamente un término, por ejemplo ‘*M*’, mediante una definición que no proporciona un criterio observacional finito para su aplicación, no excluye la posibilidad de que ‘*M*’ sea de hecho coextenso respecto de algún predicado observacional o de alguna oración compuesta por tales predicados que sea función de verdad, por ejemplo, ‘O_m’. Y si éste fuera el caso, ‘O_m’ podría entonces redefinir a ‘*M*’ y se habría proporcionado así un criterio de aplicación observacional finito.

Pero, aceptando ciertas suposiciones plausibles concernientes al vocabulario observacional, se puede probar que no todos los términos científicos pueden definirse de modo de proporcionar criterios de aplicación fini-

[38] (Añadida en 1964.) Este párrafo y algunos que lo siguen han sido modificados para corregir un enunciado erróneo de la versión original de este ensayo: que sólo las oraciones moleculares son a la vez verificables y refutables.

tos. Supondremos que el vocabulario observacional es finito. Puede contener nombres individuales que designan ciertos objetos observables, términos de predicado de primer orden con cualquier número finito de lugares que representan propiedades y relaciones de objetos observables y también functores, o sea, términos que expresan aspectos cuantitativos de los objetos observables, tales como peso en gramos, volúmenes en centímetros cúbicos, edad en días. Sin embargo, haremos de cuenta que cada uno de los functores puede asumir sólo un número finito de valores diferentes; ya que, por ejemplo, sólo se puede averiguar y distinguir mediante la observación directa un número finito de pesos diferentes.

En contraste con estos functores del vocabulario observacional, el vocabulario teórico de la física, por ejemplo, contiene un gran número de functores cuyos valores admisibles abarcan el dominio de todos los números reales o el de los números reales dentro de un cierto intervalo. Así, por ejemplo, la distancia entre dos puntos puede tener teóricamente un valor cualquiera que no sea negativo. Ahora bien, una definición del tipo requerido para un functor teórico tendría que determinar, para cada uno de sus valores admisibles, un criterio de aplicación observacional finito. Así, en el caso del functor teórico 'longitud' tendría que haber una condición necesaria y suficiente, en forma de criterio de aplicación observacional finito, para cada uno de los enunciados infinitamente numerosos de la forma 'La distancia, en centímetros, entre los puntos x e y es r', o abreviando, '$l(x, y) = r$', donde r es un número real.

En consecuencia, tendríamos que determinar para cada valor de 'r' una configuración correspondiente de observable que se pudiera establecer de modo finito. Pero esto es imposible ya que, si se tienen en cuenta los límites de la discriminación en la observación directa, sólo se podrá averiguar y distinguir un número finito, aunque muy grande, de configuraciones observables de modo finito.

Sin embargo, si no requerimos un criterio de aplicación observacional finito para cada valor admisible de un factor teórico, quizá resulten disponibles una infinidad de valores diferentes.[39] Consideremos, por ejemplo, el functor 'el número de células contenido en el organismo y'. Si se admite 'x está contenida en y' como expresiones observacionales, es posible entonces dar un criterio de aplicabilidad distinto, en términos de observables, para cada uno de los infinitos valores 1, 2, 3, ... que ese functor puede asumir teóricamente.[40] Esto se puede hacer mediante el análisis de los

[39] Agradezco a Herbert Bohnert quien, en una conversación que mantuvimos, me proporcionó el estímulo para desarrollar las ideas que he esbozado aquí acerca de la posibilidad de definir los functores con infinitos valores admisibles. El doctor Bohnert observó en esa ocasión que la definición explícita de dichos functores en términos de un vocabulario observacional debería ser posible siguiendo las líneas indicadas por la teoría de los números naturales y de los números reales, propuesta por Frege y Russell.

[40] Si se objetara que 'célula' y 'organismo' son términos teóricos en vez de observacionales, se los podría sustituir, sin afectar lo esencial del argumento, por términos cuyo carácter observacional es menos controvertible, tales como 'bolita' y 'bolsa para guardar bolitas', por ejemplo.

números cardinales propuesto por Frege y Russell. Para $n = 1$, por ejemplo, la condición necesaria y suficiente es la que se expone a continuación:

(7.10) $(\exists u)(v)[y$ es un organismo $\cdot$ $[(v$ es una célula $\cdot$ v está contenida en $y) \equiv (v = u)]]$

Así, se extiende mucho el alcance de la definición explícita en términos de observables, aun en el cálculo funcional de primer orden, si se permiten cuantitificadores en el definiens. Y si se aprueban medios lógicos más fuertes, se puede ampliar aun más el campo. Por ejemplo, se puede definir explícitamente el functor 'el número de células contenidas en y' mediante la expresión singular:

(7.11) $\hat{\alpha}$ $[\alpha$ sim. $\hat{x}$ $(x$ es una célula $\cdot$ x está contenida en $y)]$.

Aquí el acento circunflejo es el símbolo de la abstracción de clase; y 'sim', el símbolo de similitud de clases (en el sentido de correspondencia biunívoca de sus elementos).

Hasta aquí, hemos examinado sólo factores cuyos valores son enteros. ¿Se puede, en forma similar, definir a functores con valores racionales y aun irracionales en términos de observables? Consideremos, por ejemplo, el functor teórico 'longitud en centímetros'. ¿Es posible expresar, en términos observacionales, una condición necesaria y suficiente para:

(7.12) $l(x, y) = r$

aplicable a todo valor de r que no sea negativo? Podríamos intentar elaborar una definición adecuada que correspondiera al método fundamental de medición de longitudes por medio de varas rígidas. Y en verdad, si nuestro vocabulario observacional contiene un nombre para el metro patrón, y más aún, los términos (puramente cualitativos) que se requieren para describir el procedimiento de medición fundamental, es posible enunciar, para un determinado valor racional o irracional cualquiera de r, una condición necesaria y suficiente para (7.12). Sin embargo, el definiens, en la mayor parte de los casos, tendrá abundantes cuantificadores individuales, cuantificadores de clases y relaciones de diversos tipos; así, estará lejos de proporcionar un criterio de aplicación observacional finito. Indicaremos en un breve esbozo cómo se pueden obtener esas definiciones. Se escribirán en bastardilla las expresiones que, según hemos supuesto, pertenecen al vocabulario observacional.

Primero, se dirá que *el segmento determinado por los puntos x, y* tiene una longitud de cien centímetros si *es congruente* con (o sea, si se puede hacer que coincida con *el segmento marcado en el metro patrón*. Consideremos a continuación el criterio observacional para un valor racional de longitud, por ejemplo, para $l(x, y) = 0{,}25$. Se lo podrá enunciar de la siguiente forma: hay cuatro *segmentos*, cada uno *marcado en un cuerpo rígido*, tales que: 1) los cuatro son *congruentes* entre sí, 2) su *suma* (o sea, el segmento que se obtiene colocándolos extremo contra extremo a lo largo de una línea recta) es *congruente* con el *segmento marcado en el metro patrón*; 3) cualquiera de los cuatro segmentos es *congruente con el segmento determinado por los puntos x, y*. Análogamente, se puede formu-

lar un definiens observacional explícito para cualquier otro valor de n que sea múltiplo de 100 y, en consecuencia, para cualquier valor racional de n.

Segundo, la consideración de que se puede concebir un número irracional como límite de una secuencia de números racionales permite la siguiente condición necesaria y suficiente para $l(x, y) = r$, donde r es irracional: *el segmento determinado por los puntos* x,y contiene una sucesión infinita de *puntos* x_1, x_2, $x_3 \ldots$ tales que: 1) x_1 está *entre* x e y, x_2, *entre* x_1 e y, y así sucesivamente; 2) dado un *segmento* S de longitud racional, hay un *punto* en la sucesión, por ejemplo, x_n tal que los *segmentos determinados por* x_n e y, x_{n+1} e y, etc., son *más cortos que* S; 3) las longitudes de los *segmentos determinados por* x y x_1, x y x_2, y así sucesivamente, forman una secuencia de números racionales con límites r.

Finalmente, se puede usar la idea subyacente en la definición anterior para formular un definiens explícito de la expresión '$l(x,y)$' de modo tal que su dominio de valores sea el conjunto de todos los números que no son negativos.

Las definiciones del tipo que aquí se ha esbozado sólo pueden alcanzarse al costo de usar un fuerte aparato lógico: una lógica de conjuntos adecuada para el desarrollo de la teoría de los números reales.[41] Ese precio parecerá demasiado elevado a los nominalistas que sostienen que muchos de los conceptos y principios lógicos requeridos aquí, empezando por el concepto general de conjunto, son intrínsecamente oscuros y, por lo tanto, no se los debería usar en una pretendida elucidación de los significados de términos científicos. Pero éste no es el lugar para exponer las restricciones de los nominalistas; y además, la construcción de definiciones explícitas en términos de observables para un conjunto de expresiones científicas teóricas sería considerada sin duda como un valioso avance.

Se podría suscitar otra objeción contra el procedimiento esbozado para formular definiciones: considera en forma esquemática y en exceso simplificada a la medición fundamental de longitud, y es bastante liberal para interpretar como observacionales ciertos términos necesarios en el *definiens*, tales como 'cuerpo rígido' y 'punto'. Esto es cierto. Incluyendo al término punto en el vocabulario observacional, por ejemplo, consideramos a los puntos como objetos físicos directamente observables; pero nuestro criterio observacional, para dos puntos x e y que determinaban un segmento de longitud irracional, requería que hubiera una secuencia infinita de otros puntos entre x e y. Esta condición no la satisfacen nunca los "puntos" observables en la forma de pequeños objetos físicos, o señales sobre cuerpos rígidos, que se usan en la medición fundamental de longitud. Como consecuencia, la ejecución real de la medición fundamental, tal como la representó la definición anterior no dará jamás un valor irracional para la longitud de un segmento. Lo que no quiere decir que no se haya asig-

41 El argumento puede extenderse con facilidad a los functores que toman como valores a números complejos o vectores de cualquier número de componentes. Nuestro razonamiento se ha apoyado esencialmente sobre el método de definir los diversos tipos de números (enteros, racionales, irracionales, complejos, etcétera) en términos de los conceptos de la lógica de clases, tal como lo proponen Frege y Russell. Para un esbozo detallado del procedimiento, véase Russell (1919); se podrán encontrar exposiciones técnicas más completas en otros trabajos de lógica simbólica.

nado significado alguno a las longitudes irracionales; nuestro esbozo de definición muestra, por el contrario, que se puede formular un significado en términos observacionales para la asignación de cualquier valor irracional determinado que se confiera a la longitud de un segmento de línea del mundo físico, así como a la función 'longitud en centímetros' en general.

Sin embargo, el concepto de longitud así definido no es adecuado para una teoría física que incorpore a la geometría, por ejemplo, en su forma euclídea. Ya que esta última requiere que la longitud de ciertos segmentos que son accesibles a la medición directa —tales como la diagonal de un cuadrado cuyos lados tienen una longitud de 100 centímetros— tengan un valor de longitud irracional; y los enunciados al respecto resultarán siempre falsos si se considera al criterio que se acaba de exponer como estrictamente definitorio de la longitud; puesto que ese procedimiento, como ya lo destacamos, dará siempre un valor racional para la longitud de un segmento dado.

El argumento precedente acerca de los términos cuantitativos (representados por functores) muestra, en resumen, lo siguiente: el hecho de que el conjunto de valores admisibles de un functor teórico sea infinito excluye que se lo defina explícitamente mediante un vocabulario finito que sólo contenga términos cualitativos que son de carácter observacional si se adoptan criterios liberales dentro de lo razonable. Sin embargo, el argumento no muestra que se disponga de una definición semejante para todo functor requerido por la ciencia (aun nuestra definición ejemplificadora de 'longitud' resultó insatisfactoria para las necesidades de la física teórica); y en verdad, como se ha señalado antes en esta sección, no se puede esperar una prueba general a este respecto.

Varios autores han adoptado la posición de que aun cuando en principio se pudieran eludir los términos teóricos y sustituirlos por observacionales, proceder así sería prácticamente imposible o, lo que es más serio, metodológicamente desventajoso o aun obstructivo.

Veamos, por ejemplo, la respuesta que dieron Tolman y Spence al problema del que se ocupó Hull, ya mencionado en la sección 5 de este ensayo: si las variables teóricas intervinientes pueden establecer una relación segura entre condiciones observables antecedentes y consecuentes, ¿por qué no usar una conexión funcional que relacione directamente antecedentes y consecuentes? Spence aduce como argumento la consideración siguiente, también sugerida por Tolman:[42] la función matemática que se requiere para expresar la conexión serán tan compleja que resultará humanamente imposible concebirla de primera instancia: podremos llegar a ella sólo desmembrándola en una sucesión de conexiones funcionales más simples, mediatizadas por variables intermedias. Este argumento, pues, atribuye a la introducción de entidades teóricas inobservables un papel práctico importante en el descubrimiento de interdependencias entre observables y, cabe presumirlo, también en la realización efectiva de los cálculos que se requieren en la explicación o la predicción de sucesos sobre la base de tales interdependencias.

[42] Véase Tolman (1936), en la reimpresión de Marx (1951), pág. 89, y Spence (1944), pág. 65.

Hull atribuye una importante función metodológica a entidades hipotéticas en su ensayo acerca de las variables intervinientes en la teoría molar de la conducta.[43] Supongamos que, con el fin de explicar o predecir la respuesta de un sujeto en una situación dada, atribuimos al sujeto en el instante t_1 de su respuesta, cierta fuerza de hábito que tiene el status de entidad hipotética. En la teoría de Hill esa fuerza es "simplemente una representación cuantitativa de los efectos observables previos tales como estímulos observables recibidos en situaciones de aprendizaje muy anteriores. En consecuencia, si evitáramos la referencia a una entidad hipotética —la fuerza del hábito— relacionando en forma directa la respuesta observable que proporciona el sujeto en t_1 con el estímulo observable recibido antes, estaríamos invocando como determinantes causales de la respuesta ciertos acontecimientos observables que en el momento de la respuesta ya habrían dejado de existir desde mucho tiempo atrás. Y Hull rechaza esta noción de acción causal a través de una distancia temporal: "Es difícil creer que un acontecimiento tal como un estímulo en una situación de aprendizaje pueda ser causalmente activo mucho después de haber dejado de actuar sobre los receptores. Estoy totalmente de acuerdo con Lewin en que todos los factores que se alegan como causalmente influyentes en la determinación de cualquier acontecimiento deben existir en el instante de esa acción causal".[44] La referencia a la fuerza del hábito en el sujeto en el instante t_1 de su respuesta permite una explicación acorde con este principio.

Pese a que la parte final del pasaje que se acaba de citar parece bastante metafísica, el alcance básico del argumento de Hull es metodológico. Atribuye a la suposición de entidades hipotéticas explicativas un éxito que Feigl por su parte describe en otro contexto: "El carácter histórico y discontinuo (acción a distancia espacial y/o temporal) de la descripción limitada a términos fenomenistas desaparece y la reemplaza una formulación continua (contigua) en el espacio y en el tiempo y nomológicamente coherente en el nivel de la construcción hipotética".[45] Tales teorías que suponen la continuidad espaciotemporal parecen recomendables al menos por dos razones: primero, poseen cierta simplicidad formal que es difícil caracterizar hoy en términos precisos, pero que se refleja, por ejemplo, en la posibilidad de usar la poderosa y elegante maquinaria matemática del cálculo diferencial e integral para la deducción de conexiones explicativas entre los sucesos particulares a partir de los postulados de la teoría. Segundo, como se mencionó en la sección 3, el desarrollo pasado de la ciencia empírica parece mostrar que es posible encontrar que los principios explicativos y predictivos que afirman conexiones discontinuas entre acontecimientos observables (separados en el espacio y en el tiempo) tienen un alcance limitado y excepciones de diversos tipos. Con frecuencia el uso de teorías que contienen entidades hipotéticas permite dar cuenta de esas excepciones mediante suposiciones adecuadas concernientes a dichas entidades hipotéticas.

43 Hull (1943).
44 Hull (1943), pág. 285.
45 Feigl (1950), pág. 40.

Otro argumento más general es el que ha desarrollado Braithwaite, quien reconoce en Ramsey al autor de su principio básico.[46] La consideración principal de Braithwaite es que "'sólo se puede definir a los términos teóricos por medio de propiedades observables a condición de que no se pueda adaptar adecuadamente la teoría para aplicarla a nuevas situaciones".[47] Elabora esta idea refiriéndola a un modelo en miniatura de una teoría interpretada que ha formulado con precisión. Sin entrar en detalles de ese modelo —lo que requerirían aquí una digresión demasiado larga— me parece que se puede ilustrar el argumento de Braithwaite con el siguiente ejemplo: supongamos que se interprete el término 'temperatura' en una cierta etapa de la investigación científica sólo con referencia a las lecturas de un termómetro de mercurio. Si se considera a este criterio observacional como una interpretación parcial y nada más (como una condición suficiente pero no necesaria), se deja abierta la posibilidad de añadir nuevas interpretaciones parciales con referencia a otras sustancias termométricas que se puedan usar por encima del punto de ebullición o por debajo del punto de congelación del mercurio; esto permite un gran aumento del dominio de aplicación de leyes tales como las que relacionan la temperatura de una vara metálica con su longitud o su resistencia eléctrica o la temperatura de un gas con su presión o su volumen. Sin embargo, si se da al criterio original la categoría de un definiens completo, la teoría no puede extenderse de ese modo: antes bien, se debe abandonar la definición original en favor de otra que es incompatible con la primera.[48]

El concepto de inteligencia se presta para un argumento similar: si se otorga la categoría de definiciones completas a los criterios de los tests, que presuponen por parte del sujeto la capacidad de leer o al menos de usar el lenguaje ampliamente, se suscitan dificultades de ese tipo al extender el concepto y la teoría correspondiente a niños muy pequeños y animales.

Difícilmente podrá decirse que el argumento que se ha esbozado aquí prueba que "una teoría de la que se espera pueda extenderse en el futuro para explicar más generalizaciones de las que estaba destinada a explicar en su origen, debe permitir más libertad a sus términos teóricos de la que tendrían si fueran construcciones lógicas obtenidas a partir de entidades observables" [49] (definidas así en términos de dichas entidades) —y esto era lo que el argumento pretendía probar—. Porque es evidente que el procedimiento de extender una teoría al costo de cambiar las definiciones de algunos términos no es defectuoso desde el punto de vista lógico; tampoco se puede decir que sea difícil o inconveniente para el científico, ya que el problema en cuestión interesa más bien al metodólogo o al lógico, que buscan dar una clara "elucidación" o "reconstrucción lógica" de los cambios que aparecen al extender una teoría determinada. En el caso que trata Braithwaite, por ejemplo, se puede proceder de dos modos diferentes:

[46] Véase el ensayo "*Theories*" en Ramsey (1931).
[47] Braithwaite (1953), pág. 76.
[48] También Carnap hace esta observación (1936-1937, sección 7) al exponer las ventajas de las oraciones reductivas sobre las definiciones. Feigl, a su vez, argumenta de modo semejante en su ensayo de 1951, en el que ilustra el principio general mediante ejemplos de la física y la psicología.
[49] Braithwaite (1953), pág. 76.

ya sea por adiciones a la interpretación parcial original, o por un cambio total en la definición de algunas expresiones teóricas. Si se sostuviera que este último método no constituye una extensión de la teoría original sino la transición a una nueva, se suscitaría una cuestión terminológica más que una objeción metodológica.

A pesar de que el argumento anterior contra la definición no tiene el peso sistemático que se intentó darle, pone de relieve un aspecto heurístico importante del teorizar científico: cuando un científico introduce entidades teóricas, tales como corrientes eléctricas, campos magnéticos, valencias químicas o mecanismos subconscientes, procura que sirvan como factores explicativos dotados de una existencia independiente respecto de los síntomas observables por los cuales se manifiestan; o, para decirlo en forma sucinta, cualesquiera que sean los criterios observacionales de aplicación que el científico pueda proporcionar, lo que tienen por objeto es simplemente describir síntomas o indicaciones de la presencia de la entidad en cuestión y no dar una caracterización exhaustiva de ella. El científico desea en realidad dejar abierta la posibilidad de añadir a su teoría nuevos enunciados que incluyan sus términos teóricos; enunciados que pueden permitir nuevas relaciones interpretativas entre términos observacionales y teóricos; y aun considerará a éstas como nuevas suposiciones acerca de las mismas entidades hipotéticas a las que hacían referencia los términos teóricos antes de la extensión. Este modo de considerar a los términos teóricos parece tener un valor heurístico definido. Estimular la invención y el uso de conceptos con gran poder explicativo para los cuales se pueden indicar en un primer momento sólo algunos enlaces con la experiencia, pero que son fructíferos en tanto sugieren nuevas líneas de investigación que pueden a su vez conducir a relaciones adicionales con los datos de la observación directa.[50]

El examen que hemos hecho en la presente sección no ha conducido a ningún argumento concluyente en pro o en contra de la posibilidad de definir explícitamente todos los términos teóricos de la ciencia empírica por medio de un vocabulario puramente observacional, y de hecho hemos encontrado fuertes razones para dudar de que cualquier argumento pueda decidir esta cuestión de una vez por todas.

En cuanto a los términos teóricos en uso, no es imposible en la actualidad formular definiens observacionales para todos ellos que los hagan en principio innecesarios. En efecto, actualmente se usan muchos términos teóricos en la ciencia sobre la base de una interpretación empírica que es sólo parcial y este uso parece ofrecer claras ventajas heurísticas, como lo hemos señalado.

En vista de la importancia que se atribuye así a la idea de interpretación parcial, consideraremos ahora qué versión formal se puede dar de ella, y después plantearemos si es posible aplicar —y en caso afirmativo, en qué sentido lo es— el veredicto de prescindibilidad, tal como lo procla-

[50] Podrá encontrarse una concisa sinopsis de los diversos argumentos a favor del empleo de "construcciones hipotéticas" en Feigl (1950, págs. 38-41). Algunos de los aspectos del "realismo semántico" con respecto a los términos teóricos que Feigl presenta en ese mismo artículo serán expuestos en la décima sección de este ensayo.

mó "el dilema del teórico" a los términos teóricos que han sido sólo parcialmente interpretados y a los que, por lo tanto, no se puede renunciar simplemente en virtud de la definición.

8. *Sistemas interpretativos*

La teoría de Carnap de las proposiciones reductivas es el primer estudio sistemático acerca de la lógica de la definición parcial. La introducción de un término mediante una cadena de oraciones reductivas difiere del uso de una cadena de definiciones en dos aspectos significativos. Primero, se especifica el significado del término sólo parcialmente y no se proporciona así un modo de eliminar al término de todos los contextos en que pueda aparecer. Segundo, como regla, no equivale sólo a una convención notacional, sino que comprende afirmaciones empíricas. Por ejemplo, si se introduce el término '*Q*' mediante las dos oraciones reductivas:

(8.1) $C_1x \supset (Qx \equiv E_1x)$

(8.2) $C_2x \supset (Qx \equiv E_2x)$

por implicación se afirmará la siguiente ley:

(8.3 $(x)[(C_1x \cdot E_1x) \supset (C_2x \supset E_2x)]$

o sea, para decirlo en forma burda: cualquier objeto que muestre una respuesta positiva en la primera condición de prueba mostrará también una respuesta positiva sometido a la segunda condición de prueba. Así, una cadena de oraciones reductivas para un término dado combina normalmente dos funciones del lenguaje que se consideran a menudo como rigurosamente distintas: la asignación estipulativa de significado y la afirmación o descripción del hecho empírico.

Como vimos antes, las oraciones reductivas se adecuan muy bien a la formulación de criterios operacionales de aplicación en tanto definiciones parciales. Pero están sujetas a limitaciones bastante serias en lo que respecta a la forma lógica y así, no parecen suficientes para proporcionar un esquema general satisfactorio para la interpretación parcial de los términos teóricos.[51] La concepción de Campbell de la teoría física, según la cual ésta consiste en una "hipótesis", representada por un conjunto de oraciones en términos teóricos y un "diccionario" que relaciona a estos últimos con los conceptos de la física experimental (que deben estar interconectados por leyes empíricas) sugiere una visión más amplia que la anterior respecto de la interpretación.[52] En contraste con la concepción habitual de diccionario, el de Campbell no ha de contener definiciones de los términos teóricos, sino enunciados para que una orientación teórica de un cierto tipo sea verdadera si y sólo si la correspondiente oración empírica de un tipo determinado es verdadera. Así, el diccionario proporciona, más que defi-

[51] Esto lo ha señalado el mismo Carnap; véase por ejemplo (1956).
[52] Véase Campbell (1920, cap. VI). Las partes importantes de este capítulo están reimpresas en Feigl y Brodbeck (1953).

niciones, reglas de traducción y aun reglas parciales, porque nadie pretende que se deba especificar una traducción para cada enunciado teórico o para cada enunciado empírico.

Este último rasgo concuerda, por ejemplo, con la consideración de que un particular macroestado observable de un determinado sistema físico puede corresponder a un gran número de microestados distinguibles desde el punto de vista teórico: de modo tal que, para una oración teórica que describa uno de aquellos microestados, la oración que ha de describir el macroestado correspondiente no expresa una condición necesaria y suficiente y por lo tanto no proporciona traducción alguna.[53]

Es evidente que los enunciados del diccionario de Campbell no tienen el carácter de oraciones reductivas; pueden formularse, sin embargo, como bicondicionales en los que una oración en términos teóricos se conecta con otra en términos observacionales mediante la cláusula "si y sólo si".

En otros contextos no parecen adecuadas ni las oraciones reductivas ni esas bicondicionales. Porque en general, la presencia de una entidad hipotética H, tal como un cierto tipo de campo eléctrico, tendrá síntomas observables sólo si se satisfacen ciertas condiciones observacionales, O_1, tales como la presencia de detectores adecuados, que tendrán que mostrar respuestas observables, O_2. Una oración que enunciara ese tipo de criterio tendría el carácter de una oración reductiva generalizada; se la podría formular del siguiente modo:

(8.4) $O_1 \supset (H \supset O_2)$

donde 'O_1' y 'O_2' son oraciones —posiblemente bastante complejas— en términos de observables, y 'H' es una oración expresada en términos teóricos.

Pero no hay razones de peso para limitar los enunciados interpretativos a los tres tipos considerados aquí. Para obtener un concepto general de la interpretación parcial, admitiremos ahora como enunciados interpretativos a toda oración, cualquiera sea su forma lógica, que contenga términos teóricos y observacionales. Si suponemos que los enunciados teóricos y observacionales de la ciencia empírica se formulan dentro de un marco lógico determinado, esa idea puede enunciarse más precisa y explícitamente como sigue:

(8.5) Sea T una teoría que *se caracteriza* por el conjunto de postulados en términos de un *vocabulario teórico* finito V_T y sea V_B un segundo conjunto de términos extralógicos, al que se llamará *vocabulario básico,* que no comparte ningún término con V_T; entenderemos por *sistema interpretativo* para T con base V_B un conjunto J de oraciones tal que: 1) es finito, 2) es lógicamente compatible con T, 3) no contiene términos extralógicos que no pertenezcan a V_T o a V_B, 4) contiene esencialmente a todo elemento de V_T y de

[53] Sin embargo, esto no muestra que no exista la posibilidad de una condición necesaria y suficiente en términos observacionales para la oración teórica: el problema de probar o refutar este argumento está sujeto a dificultades análogas a las expuestas en la sección 7 respecto de la definibilidad.

V_B o sea, no es lógicamente equivalente a ningún conjunto de oraciones en las que no aparezca para nada algún término de V_T o de V_B.[54]

Al aplicar el concepto que se ha definido aquí al análisis de las teorías científicas naturalmente tendremos que suponer que los términos de que consta V_B han sido ya comprendidos. Pueden ser términos observacionales en el sentido algo vago que se ha explicado antes, pero no necesitamos insistir sobre eso. Se puede adoptar el punto de vista, por ejemplo, de que ciertos términos disposicionales, tales como 'maleable', 'elástico', 'hambriento', 'cansado' no son estrictamente observacionales, y no se sabe que puedan definirse explícitamente por medio de términos observacionales; aun así, se puede considerar que se comprende bien a esos términos en el sentido de que diferentes observadores competentes coinciden en alto grado al usarlos. En este caso sería bastante razonable usar esos términos para interpretar una teoría dada, o sea admitirlos en V_B.

La concepción que tiene Campbell de la función de su "diccionario" ilustra muy bien esa posibilidad y muestra que se acerca más al procedimiento científico real. Campbell estipula que la interpretación que proporciona el diccionario debe constar de lo que él llama "conceptos", tales como los términos 'temperatura', 'resistencia eléctrica', 'plata', 'hierro' en el sentido en que se los usa en la física y la química experimental. Es difícil que éstos se puedan considerar observacionales en sentido estricto ya que se los ha concebido específicamente para representar grupos de leyes empíricas. "Así, cualquier cosa que digamos de la resistencia eléctrica, estaremos suponiendo que la ley de Ohm es verdadera; los cuerpos para los cuales dicha ley no es verdadera, los gases, por ejemplo, no tienen resistencia eléctrica".[55] Pero, aun cuando no se deseara calificar a esos términos de observacionales, se podría considerar que se los comprende bien y que los experimentadores científicos coinciden en alto grado al usarlos, y se los podría admitir así en V_B.

Los sistemas interpretativos, tal como se los acaba de definir, incluyen como casos especiales a todos los tipos de interpretación que consideramos antes: la interpretación por medio de definiciones explícitas para todos los términos teóricos, por medio de cadenas de oraciones reductivas, por medio de enunciados traductorios bicondicionales en el sentido del diccionario de Campbell, y por medio de oraciones reductivas generalizadas de la forma (8.4); pero, por supuesto, admiten también enunciados interpretativos de una gran variedad de formas diferentes de las expuestas.

[54] La noción intuitiva de interpretación, así como la concepción que pone de manifiesto la idea de Campbell de un diccionario interpretativo, parecería tener necesidad de la siguiente condición adicional: (v) Cada oración de J contiene esencialmente términos de V_T así como términos de V_B. Sin embargo, este requisito no introduce nuevas restricciones al concepto de sistema interpretativo, ya que cualquier sistema J que satisfaga las condiciones (i) a (iv) puede enunciarse en una forma equivalente tal que satisfaga asimismo a (v). Para este fin, basta reemplazar a las oraciones miembros de J por su conjunción, lo que permite un sistema interpretativo lógicamente equivalente que contiene sólo una oración y que satisface (v) puesto que J satisface (iv).

[55] Campbell (1920, pág. 43).

Los sistemas interpretativos tienen el mismo par de características que distinguían a las cadenas de oraciones reductivas de las cadenas de definiciones. Primero, un sistema interpretativo normalmente realiza sólo una interpretación parcial de los términos de V_T o sea, no establece (mediante un enunciado explícito o una implicación lógica) una condición necesaria y suficiente de aplicación en términos de V_B para todo término de V_T. Segundo, un sistema interpretativo no tendrá, por lo común, un carácter puramente estipulativo —lo mismo ocurre con una cadena de oraciones reducidas para un término teórico dado— sino que implicará ciertos enunciados en términos de V_B; solamente que no son verdades lógicas sino que, según la concepción de V_B como vocabulario formado por términos empíricos previamente comprendidos, expresan afirmaciones empíricas. También aquí encontramos una combinación de los usos descriptivo y estimulativo del lenguaje.

Pero, para pasar a un tercer punto de comparación, un sistema interpretativo no necesita proporcionar una interpretación —ya sea completa o incompleta— para cada término de V_T en forma individual. A este respecto, difiere de un conjunto de definiciones que especifica para cada término una condición necesaria y suficiente, y de un conjunto de oraciones reductivas, que proporciona para cada término una condición necesaria y una suficiente, por lo común diferentes entre sí. Es posible que un sistema interpretativo no proporcione ninguna condición suficiente, o ni siquiera una de las dos, para varios o aun para todos los términos de V_T; en cambio, podría especificar, mediante enunciados explícitos o mediante implicaciones lógicas, condiciones necesarias o suficientes en términos de V_B, por ejemplo, a la manera del diccionario de Campbell.

Por lo tanto, cuando se interpreta una teoría *T* mediante un sistema interpretativo *J*, en general no se puede prescindir de los términos teóricos en el sentido estricto de eliminarlos de todos los contextos en favor de expresiones que los definen en términos de V_B. Tampoco es posible, por lo general, prescindir de ellos en el sentido de que *J* proporcione para toda proposición *H* que pueda formarse por medio de V_T una "traducción" en términos de V_B, o sea, una proposición *O* en términos de V_B tal que el bicondicional $H \equiv O$ [56] sea lógicamente deducible de *J*.

¿Son pues, los términos teóricos indispensables en esta amplia concepción de la interpretación de modo tal que "la paradoja del teorizar", tal como se la formuló en la sección 5, ya no se aplica a ellos? Consideraremos esta cuestión en la sección siguiente.

9. *Posibilidad de un reemplazo funcional de los términos teóricos*

La función sistematizadora de una teoría *T*, tal como la interpreta un sistema interpretativo *J* consistirá en permitir inferir a partir de determi-

[56] Aquí, y en algunas ocasiones subsiguientes en las que no hay peligro de error, se usan los conectivos lógicos en forma automática; la expresión 'H ≡ O', por ejemplo, representa a la oración que se obtiene colocando la triple barra que es símbolo de 'si y sólo si' entre las oraciones a las que se ha designado 'H' y 'O'.

nados "datos" en términos de V_B, ciertos enunciados (predictivos, por ejemplo) en términos de V_B. Si O_1 es el enunciado que expresa los datos y O_2, el enunciado inferior, la conexión puede simbolizarse así:

(9.1) $(O_1 \cdot T \cdot J) \rightarrow O_2$

Aquí, como en contextos semejantes que aparecerán más adelante, 'T' representa el conjunto de los postulados de la teoría en cuestión; la flecha representa la implicación deductiva.

Ahora bien, (9.1) es verdadera si y sólo si $T \cdot J$ implica la oración $O_1 \supset O_2$; de modo tal que (9.1) es equivalente a:

(9.2) $(T \cdot J) \rightarrow (O_1 \supset O_2)$

Sea cual fuere la relación establecida entre las transiciones de V_B, resulta claro que se logra mediante T en conjunción con J. Por lo tanto, será conveniente considerar a los postulados de T junto con las oraciones de J como los postulados de un sistema deductivo T', al que llamaremos *teoría interpretada.* Su vocabulario $V_{T'}$ será la suma de V_T y V_B.

Lo que se observó en relación con (9.1) y (9.2) podemos ahora exponerlo nuevamente de este modo. Si una teoría interpretada T' establece una transición deductiva de O_1 *a* O_2, o sea, si

(9.3) $(O_1 \cdot T') \rightarrow O_2$

entonces

(9.4) $T' \rightarrow (O_1 \supset O_2)$

e inversamente, donde T' es el conjunto de postulados de la teoría interpretada.

Ahora se puede mostrar fácilmente que una teoría interpretada T' establece las mismas conexiones deductivas entre oraciones de V_B que el conjunto de todos los teoremas de T' que son expresables en términos de V_B solamente; llamaremos a este conjunto *teoremas-V_B* o *consecuencias-V_B* de T' y lo designaremos como '$O_{T'}$'. Esto significa que, para todos los propósitos de la sistematización deductiva, T' es, como diremos ahora en adelante, *funcionalmente equivalente* al conjunto O_T que no contiene ningún solo término teórico.

La prueba es la siguiente: la transición deductiva representada en (9.3) de O_1 a O_2 puede llevarse a cabo también usando en lugar de T', la oración $O_1 \supset O_2$ que en virtud de (9.4) pertenece a $O_{T'}$; ya que tenemos por *modus ponens,*

(9.5) $[O_1 \cdot (O_1 \supset O_2)] \rightarrow O_2$

Y puesto que $O_{T'}$ contiene con seguridad todas las oraciones de V_B de la forma $O_1 \supset O_2$ que implica T', el conjunto $O_{T'}$ basta para realizar todas las sistematizaciones deductivas que se pueden obtener por medio de T'. Por otra parte $O_{T'}$ no es más fuerte a este respecto que T'; ya que $O_{T'}$, permite la transición deductiva de O_1 a O_2 sólo si implica a $O_1 \supset O_2$, lo

que significa, en vista de la equivalencia de (9.4) con (9.3) que T' permitirá la transición deductiva de O_1 a O_2.

Así, *la sistematización deductiva que una teoría interpretada T' lleva a cabo entre oraciones expresadas en términos del vocabulario básico V_B es la misma que logra el conjunto $O_{T'}$, de los enunciados (teoremas) de T' que pueden expresarse en términos de V_B solamente.* En este sentido, se puede prescindir de los términos usados en T.

Pero por lo general $O_{T'}$ es un conjunto infinito de enunciados que resulta difícil de manejar; se suscita por lo tanto la cuestión de saber si hay algún método aplicable que confiriendo a O_T la forma de un sistema axiomatizado teórico T'_B, formulado en términos de V_B solamente, lo torne más manejable y claro. Un teorema en lógica formal que Craig ha probado muestra que esto se puede hacer siempre y cuando T' satisfaga ciertas condiciones muy liberales y no restrictivas.[57]

[57] Este importante teorema apareció publicado por primera vez en el artículo de Craig (1953). En Craig (1956) hay una presentación menos condensada y menos técnica, con referencias explícitas aunque breves, a aplicaciones tales como la que se ha considerado aquí.

Aplicando el resultado obtenido por Craig al tema en discusión, se lo puede enunciar en forma breve de la siguiente manera: supongamos que se especificara al conjunto $V_{T'}$, de términos primitivos de T' y al conjunto de postulados de T' en forma efectiva, o sea de modo tal de proporcionar un procedimiento general que permitiera, dada una expresión, decidir en un número finito de pasos si esa expresión es un término primitivo (o un postulado) de T' o no lo es. Supongamos que $V_{T'}$ se dividiera mediante un criterio efectivo que por lo demás, podría elegirse a voluntad, en dos vocabularios mutuamente excluyentes, V_T y V_B. Finalmente supongamos que las reglas de la lógica que se usaran permitieran un método efectivo para determinar, con respecto a cualquier sucesión finita dada de expresiones, si es una deducción válida de acuerdo con esas reglas.

Existiría entonces un método general (o sea, un método aplicable en todos los casos que satisfacen las condiciones que se acaba de esbozar) para construir en forma efectiva (o sea, caracterizar en forma efectiva los postulados y las reglas de inferencia de) un nuevo sistema T'_B cuyo conjunto de primitivos es V_B y cuyos teoremas son los mismos teoremas de T' que no contiene otras constantes extralógicas que aquellas de que consta V_B.

Adviértase que el teorema nos permitirá trazar la línea divisoria entre V_T y V_B donde nos plazca, siempre y cuando el criterio que se use para realizar la división nos permita decidir en un número finito de pasos a cuál de los dos conjuntos pertenece el término. Esta condición, así como el requisito de una caracterización efectiva de $V_{T'}$, quedarán satisfechos en forma trivial si, por ejemplo, $V_{T'}$ es finito, y se especifican sus miembros al igual que los de V_B y V_T por una simple enumeración individual.

Los otros requisitos de una caracterización efectiva de los postulados y de las reglas de la lógica para T′ son tan liberales que sin duda cualquiera de las teorías científicas que se han considerado hasta ahora puede ser formalizada de modo que los satisfaga, siempre y cuando las conexiones entre términos observacionales y teóricos puedan expresarse en forma de enunciados definidos. El único caso importante que conozco de violación de esa condición es el de una teoría para la cual no se especifican reglas definidas de interpretación, sobre la base de que los criterios de aplicación de las expresiones teóricas tendrán que permanecer siempre algo vagos. Quizá sea una concepción de este tipo la que propone A. Wald al hacer la siguiente observación: "Para aplicar una teoría [científica] a los fenómenos reales, necesitamos algunas reglas para establecer la correspondencia entre los objetos idealizados de la teoría y los del mundo real. Estas reglas serán siempre algo vagas y nunca podrán formar parte de la teoría misma." Wald (1942), pág. 1.

Sin embargo, se pueden satisfacer las condiciones del teorema de Craig si la

Así, el teorema de Craig tiene una clara proyección sobre los problemas suscitados por la "paradoja del teorizar" que se enunció en la sección 5 en términos algo vagos. El teorema citado señala un camino para dar una interpretación clara y precisa de la "paradoja" a la vez que una prueba rigurosa de ella: muestra que para cualquier teoría T' que usa a la vez términos teóricos y no teóricos previamente comprendidos y dadas ciertas condiciones satisfechas con amplitud, existe un sistema teórico axiomatizado T'_B que usa sólo los términos no teóricos de T' y sigue siendo, sin embargo, funcionalmente equivalente a T' en el sentido de que realiza las mismas conexiones deductivas que T' entre las oraciones que se pueden expresar en el vocabulario no teórico.

Siendo esto así, la ciencia empírica ¿debería aprovechar este método y reemplazar todas las teorías que incluyen suposiciones acerca de entidades hipotéticas por sistemas teóricos funcionalmente equivalentes expresados sólo en términos que tuvieran referencia observacional directa o bien que fueran por lo menos comprendidas con claridad? Hay varias razones que hacen que este reemplazo sea desaconsejable teniendo en cuenta los objetivos del teorizar científico.

Para empezar, consideremos el carácter general del método de Craig. Omitiendo muchos detalles sutiles, puede describírselo como sigue: por medio de un procedimiento constructivo, Craig dispone a todos los teoremas V_B de T' en una secuencia. Esta es muy redundante pues contiene, para toda oración que aparece en ella, todos sus equivalentes lógicos (hasta donde se los pueda expresar en V_B). Craig prescribe un procedimiento para eliminar muchas de esas duplicaciones pero no todas son eliminables. La secuencia resultante contiene todavía a cada teorema V_B de T' en una por lo menos de sus formulaciones equivalentes. Finalmente, todas las oraciones de la secuencia así formada se convierten en postulados de T'_B. Así, al conjunto de teoremas V_B de T' se lo "axiomatiza" en T'_B sólo en un sentido artificial, convirtiendo a cada oración del conjunto, en alguna de sus muchas formulaciones equivalentes, en un postulado de T'_B. Normalmente, la axiomatización de un conjunto de oraciones selecciona como postulados a un pequeño subconjunto a partir del cual se puede derivar lógicamente el resto de las oraciones que se convierten en teoremas; de esta manera, la axiomatización expresa el contenido del conjunto íntegro "en una forma que es psicológica o matemáticamente más clara".[58]

---

vaguedad a la que nos hemos referido se pone de manifiesto mediante reglas definidas. Así, por ejemplo, las oraciones interpretativas para una teoría dada podrían adoptar la forma de enunciados de probabilidad estadística (posibilidad que Carnap menciona [1956], sección 5), o quizá de enunciados de probabilidad lógica (cada uno de los cuales especificaría la probabilidad lógica de alguna oración teórica relativa a una oración especificada en términos observacionales o viceversa). Cualquiera de estos procedimientos permitiría una interpretación más general que la que caracterizó la definición de un sistema interpretativo, definición propuesta en la sección 8 de este capítulo. Incluso se puede aplicar el teorema de Craig a las teorías interpretadas en este sentido más amplio.

[58] Craig (1956), pág. 49. Convendría destacar brevemente dos observaciones más que hizo Craig en los trabajos a los que nos hemos referido aquí: (i) Una teoría T' puede tener un conjunto de consecuencias V_B que no pueden axiomatizarse mediante un conjunto *finito* de postulados expresables en términos de V_B. (ii) No hay

Puesto que el método de Craig incluye a todas las oraciones que tienen que ser axiomatizadas entre los postulados de T', este último conjunto de oraciones "no consigue simplificar o proporcionar una genuina comprensión del tema", según lo declara el mismo Craig.[59]

La pérdida de simplicidad que resulta de descartar los términos teóricos de T' se refleja en la circunstancia de que el conjunto de postulados que el método de Craig proporciona a T'_B es siempre infinito. Aun en los casos en que exista realmente algún subconjunto finito de $O_{T'}$ formado por teoremas V_B de T' a partir de los cuales se puedan deducir todos los restantes, el procedimiento de Craig no proporcionará un subconjunto de ese tipo: ése es el precio que se ha de pagar por su aplicabilidad universal.

Ahora bien, hay casos en los que una infinidad de postulados puede no ser muy difícil de manejar, en particular cuando los axiomas se expecifican mediante esquemas axiomáticos,[60] o sea, por medio de estipulaciones a los efectos de que cualquier oración que tenga una entre un número finito de formas determinadas (tales como '$x = x$', por ejemplo) sea considerada como axioma. Pero la manera en que se especifican los postulados de T'_B por el método de Craig es mucho más intrincada, y el sistema obtenido sería prácticamente inmanejable; para no hablar de la pérdida de la fertilidad y de la riqueza heurísticas que resultaría de la eliminación de las hipótesis y conceptos teóricos. Para la ciencia empírica, por lo tanto, este método de prescindir de las expresiones teóricas resultaría insatisfactorio.

Hasta aquí hemos examinado la posibilidad de eliminar suposiciones y conceptos teóricos sólo en el contexto de la sistematización deductiva: consideramos que una teoría interpretada T' era sólo un vehículo para establecer transiciones deductivas entre oraciones observacionales. Sin embargo, esas teorías pueden proveer también distintos modos de sistematización inductiva (en el sentido que se esbozó en la sección 1) y un análisis de esta función dará un argumento más en contra de la eliminación de expresiones teóricas mediante el método de Craig.

A modo de ilustración usaremos un ejemplo que ha sido simplificado en forma deliberada para exhibir con más claridad los elementos esenciales: supongamos que V_T contiene el término 'fósforo blanco' o 'F' para abreviar, y que el sistema interpretativo incorporado a T' no establece para dicho término condiciones observacionales de aplicación que sean suficientes pero sí algunas condiciones que son necesarias. Supongamos que éstas son independientes entre sí en el sentido de que, aun cuando en el caso del fósforo blanco aparezcan juntas, cualquiera de ellas podrá

---

un método general que permita para toda teoría T', decidir, en forma efectiva, si sus consecuencias V_B pueden axiomatizarse o no mediante un conjunto finito de postulados.

[59] Craig (1956), pág. 49. Por supuesto este hecho no disminuye en lo mínimo la importancia y el interés que tiene el resultado obtenido por Craig como teorema de la lógica.

[60] Sobre este método que fue usado en primer lugar por Neumann, véase Carnap (1937, págs. 29-30 y pág. 96), donde se dan ulteriores referencias a la bibliografía sobre estos temas.

aparecer en otros casos en ausencia de una o de varias de las demás. Esas condiciones necesarias podrían ser las siguientes: el fósforo blanco tiene un olor semejante al del ajo, es soluble en trementina, es soluble en aceites vegetales, es soluble en éter, produce quemaduras en la piel, que se expresan así en notación simbólica:

(9.6) $(x)\ (Fx \supset Ox)$

(9.7) $(x)\ (Fx \supset Tx)$

(9.8) $(x)\ (Fx \supset Vx)$

(9.9) $(x)\ (Fx \supset Ex)$

(9.10) $(x)\ (Fx \supset Px)$

Todos los predicados que aparecen en estas oraciones excepto '*F*' pertenecen pues a V_B.

Ahora supongamos que V_T contiene sólo un término más aparte de '*F*': se inflama a 30°C de temperatura o '*I*' para abreviar, y que hay sólo una oración interpretativa para '*I*' según la cual si un objeto tiene la propiedad *I*, arderá al estar rodeado por aire en el cual un termómetro señale una marca por encima de los 30°C. Consideraremos que esta propiedad es observable y la representaremos con el predicado '*A*' en V_B. La oración interpretativa para '*I*' es, pues:

(9.11) $(x)\ (Ix \supset Ax)$

Finalmente supongamos que la parte teórica de *T'* contiene un postulado único:

(9.12) $(x)\ (Fx \supset Ix)$

según el cual el fósforo blanco se inflama a 30°C de temperatura y que las siete oraciones (9.6)-(9.12) representan el contenido total de T'.

Luego, como se ve en seguida, *T'* no tiene consecuencias en términos de V_B excepto en lo que respecta a verdades puramente lógicas; por consiguiente, *T'* permitirá una transición deductiva de una oración de V_B a otra sólo si la primera implica lógicamente a la segunda, de modo tal que no se requiera a *T'* para establecer la conexión. En otras palabras, *T'* no realiza sistematización deductiva alguna entre las oraciones de V_B. No obstante, *T'* puede desempeñar un papel esencial para establecer ciertas conexiones explicativas o predictivas de tipo inductivo entre las oraciones de V_B. Imaginemos, por ejemplo, que se ha encontrado que cierto objeto *c* posee todas las características *O*, *T*, *V*, *E*, *P*. Teniendo en cuenta las oraciones (9.6)-(9.10), según las cuales dichas características son sintomáticas de *F*, podría inferirse que *c* es fósforo blanco. Esta inferencia sería inductiva en lugar de deductiva, y su fuerza derivaría en parte de la independencia mutua que atribuimos a esos cinco síntomas observables del fósforo blanco. La oración '*Fc*' a la que se ha llegado así en forma inductiva conduce a la predicción '*Ic*' a través de (9.12); '*Ic*' a su vez,

en virtud de (9.11) permite la predicción '*Ac*'. Así, T' permite la transición de los datos observacionales '*Oc*', '*Tc*', '*Vc*', '*Ec*', '*Pc*' a la predicción observacional '*Ac*'. Pero la transición requiere un paso inductivo que consiste en aceptar a '*Fc*' sobre la base de los cinco datos citados que lo apoyan pero, por supuesto, no lo implican lógicamente.

Por otra parte el sistema T'_B obtenido por el método de Craig no se presta para este uso inductivo; de hecho todas sus oraciones son verdades lógicas y T'_B no proporciona así ninguna afirmación empírica, ya que, como se ha observado antes, todos los teoremas V_B de T' son enunciados lógicamente verdaderos.

Así, si se considera que el uso sistematizador de una teoría interpretada T' incluye procedimientos deductivos y también inductivos, no se podrá reemplazar en general a T' por el sistema correspondiente T'_B.

Ramsey sugiere otro método, intuitivamente más simple, para obtener un equivalente funcional formulado en términos observacionales de una teoría interpretada T'. En efecto, el método intenta tratar a todos los términos teóricos como variables con cuantificación existencial de manera que todas las constantes extralógicas que aparecen al formular una teoría del modo propuesto por Ramsey pertenecen al vocabulario observacional.[61] Así, la teoría interpretada determinada por nuestras fórmulas (9.6)-(9.12) podría expresarse mediante la oración siguiente, que llamaremos *la oración de Ramsey asociada con la teoría dada*:

(9.13) $(\exists\varphi)(\exists\psi)(x)[(\varphi x \supset (Ox \cdot Tx \cdot Vx \cdot Ex \cdot Px)) \cdot (\psi x \supset Ax) \cdot (\varphi x \supset \psi x)]$

Esta oración es equivalente a la expresión que se obtiene uniendo las oraciones (9.6)-(9.12) y reemplazando en todos los casos '*F*' e '*I*' por las variables 'φ' y 'ψ', respectivamente, y anteponiendo a dichas variables cuantificadores existenciales. Así (9.13) afirma que hay dos propiedades φ y ψ, no especificadas por otra parte, tales que cualquier objeto con la propiedad φ tiene también las propiedades observables *O, T, V, E, P* y cualquier objeto con la propiedad ψ tiene también la propiedad observable *A*.

Una teoría interpretada T' por supuesto, no es lógicamente equivalente a la oración asociada de Ramsey como tampoco lo es al sistema asociado T'_B de Craig; de hecho, cada uno de los dos está implicado por T' pero no implica a su vez a T'. A pesar de que la oración de Ramsey contiene sólo términos de V_B, aparte de las variables y constantes lógicas, se puede mostrar que implica las mismas oraciones V_B que T'; en consecuencia, establece las mismas transiciones deductivas entre oraciones V_B que T'. A este respecto, pues, la oración de Ramsey asociada con T' está en paridad de condiciones con el sistema T'_B de Craig que puede obtenerse de T', pero su aparato lógico es mucho más complicado que el requerido por T' o por T'_B. En nuestro ejemplo, T' y T'_B contienen variables y cuantificadores sólo con respecto a individuos (objetos físicos) en tanto que la ora-

61 Ramsey (1931), págs. 212-215, 231.

ción de Ramsey (9.13) contiene variables y cuantificadores también para propiedades de individuos; así, mientras T' y T'_B sólo requieren un cálculo funcional de primer orden, la oración de Ramsey necesita un cálculo funcional de segundo orden.

Pero esto significa que la oración de Ramsey asociada con una teoría interpretada T' evita hacer referencia a entidades hipotéticas sólo en la letra —reemplazando las constantes latinas por variables griegas— y no en el espíritu. Ya que afirma aún la existencia de ciertas entidades del tipo que postula T', sin garantizar más de lo que lo hace T', que esas entidades sean observables o al menos que se las pueda caracterizar totalmente en términos de observables. En consecuencia, las oraciones de Ramsey no proporcionan un modo satisfactorio de eludir los conceptos teóricos.

A decir verdad, el mismo Ramsey no tuvo esa pretensión. Antes bien, su interpretación de los términos teóricos como variables con cuantificación existencial parece haber sido motivada por consideraciones del siguiente tipo: si se trata a los términos teóricos como constantes que no han sido definidas por completo mediante términos observacionales ya comprendidos, las oraciones que se pueden construir con ellos no tienen el carácter de afirmaciones con significados especificados en forma total a las que se puede atribuir significativamente valor de verdad, sino que su status es comparable al de las funciones oracionales donde los términos teóricos como variables con cuantificación existencial permite una formulación de las mismas que no sólo satisface el requisito de que sean verdaderas o falsas, sino que al mismo tiempo conserva todas sus implicaciones empíricas.

Esta consideración suscita un nuevo problema que se tratará en la sección siguiente.

10. *Sobre el significado y la verdad de las teorías científicas*

He aquí el problema sugerido por el enfoque de Ramsey: si se interpreta a los términos teóricos de una teoría como constantes extralógicas para las cuales el sistema J proporciona sólo una interpretación parcial en términos del vocabulario previamente comprendido V_B, tal como se hizo en la sección 8, ¿se podrá considerar no obstante a las oraciones formadas mediante el vocabulario teórico como oraciones significativas que hacen afirmaciones definidas y que son o verdaderas o falsas?

La cuestión parecería caer bajo la jurisdicción de la semántica y, más específicamente, de la teoría semántica de la verdad. Pero no ocurre así. Lo que proporciona la teoría semántica (en ciertas condiciones) es una definición general de verdad para las proposiciones de un lenguaje dado L. Esa definición se enuncia en un metalenguaje M adecuado para L y permite la formulación de una condición de verdad necesaria y suficiente para cualquier oración P de L. Esta condición se expresa mediante la traducción de P a M.[62] Para adecuarse a su propósito, M debe obtener

[62] Véase Tarski (1944), sección 9.

por lo tanto, una traducción de cada oración de L y satisfacer otras condiciones que se especifican en la teoría semántica de la verdad. Pero si los criterios de verdad enunciados en M han de ser inteligibles, resulta claro que para empezar, todas las traducciones de enunciados de L a M deben ser consideradas significativas. En lugar de decidir la cuestión de la significatividad de los enunciados de L, la definición semántica de verdad presupone que ya se la ha resuelto.

Por razones análogas, la semántica no nos capacita para decidir si los términos teóricos de un sistema dado tienen referencia semántica o fáctica u ontológica o bien no la tienen, siendo esta característica la que para algunos autores distingue a los conceptos genuinamente teóricos de los términos teóricos auxiliares o intervinientes.[63] Los argumentos opuestos que se han formulado para debatir esta cuestión presentan una dificultad y es que no se indica con claridad qué se quiere decir cuando se atribuye una referencia ontológica a un término dado. Desde un punto de vista puramente semántico es posible atribuir referencia semántica a cualquier término del lenguaje L que se considere comprendido: el referente puede especificarse de la misma manera que la condición de verdad de una oración dada en L, por traducción a un metalenguaje apropiado. Por ejemplo, podríamos usar el castellano como metalenguaje con respecto a la terminología de Freud; diríamos entonces que 'Verdraengung' designa represión, 'Sublimierung', sublimación, y así sucesivamente. Es evidente que este tipo de información es poco esclarecedora para los que desean usar la referencia existncial como una característica distintiva de cierto tipo de términos teóricos: tampoco ayuda a los que quieren saber si se puede decir que las entidades designadas por los términos teóricos existen realmente, y en caso afirmativo, en qué sentido —cuestión a la que volveremos un poco más adelante—.

La semántica, pues, no responde a la pregunta que hicimos al comienzo de esta sección: tenemos que buscar en otra parte los criterios de significación para las expresiones teóricas.

Generalizando, podríamos calificar a una expresión teórica de inteligible o significativa si se la ha explicado en forma adecuada con términos que consideramos han sido comprendidos previamente. En nuestra exposición anterior los términos de ese tipo estaban representados por el vocabulario V_B (más los términos de la lógica). Pero ahora se suscita la cuestión: ¿qué es lo que constituye una explicación "adecuada"? No se pueden determinar normas obligatorias generales: la respuesta en última instancia la da cada uno de acuerdo con su conciencia filosófica. El puritano lógico y gnoseológico declarará inteligible sólo lo que ha sido definido en forma explícita en términos de V_B y podrá imponer nuevas restricciones, a la manera de un nominalista, por ejemplo, al aparato lógico que se puede usar para formular definiciones. Otros encontrarán que las oraciones reductivas son inteligibles y otros apoyarán hasta una interpretación tan endeble como la que lleva a cabo un sistema interpretativo. Una de las ventajas más importantes de la definición es que asegura la

[63] Sobre este punto véase, por ejemplo, MacCorquodale y Meehl (1948); Lindzey (1953); Feigl (1950) y (1950a); Hempel (1950); Rozeboom (1956).

posibilidad de formular en términos de V_B un enunciado equivalente a una oración teórica cualquiera. La interpretación parcial no garantiza esto; por consiguiente, no proporciona para cada oración expresable en términos teóricos, una condición de verdad necesaria y suficiente que puede enunciarse en términos que han sido previamente comprendidos. Indudablemente, ésta es la dificultad básica que encuentran los críticos en el método de la interpretación parcial.

En cambio, puede decirse para defender la interpretación parcial, que comprender una expresión es saber cómo usarla, y en una reconstrucción formal, el 'cómo' se expresa mediante reglas. La interpretación parcial, tal como la hemos concebido, proporciona reglas de ese tipo y éstas muestran, por ejemplo, qué oraciones expresadas sólo en términos de V_B pueden inferirse a partir de oraciones que contienen términos teóricos: y así especifican a un conjunto de oraciones V_B tales que una teoría interpretada *T'* las implica y en consecuencia, indirectamente las afirma. (Si el conjunto es vacío, la teoría no cae bajo el dominio de la ciencia empírica.) Inversamente, las reglas muestran también que las oraciones formuladas en términos teóricos pueden inferirse a partir de las oraciones V_B. Así, hay marcadas semejanzas entre nuestras oraciones teóricas y las que son inteligibles en el sentido más estrecho de ser expresables por completo en términos de V_B —circunstancia que favorece la admisión de las oraciones teóricas dentro de la clase de los enunciados significativos—.

Debería añadirse que, de adoptar esta línea de acción, tendremos que reconocer que ciertos sistemas interpretativos que con seguridad no calificaríamos de teorías científicas en potencia son, sin embargo, significativos (aunque por supuesto, no sean interesantes ni valga la pena investigarlos). Por ejemplo, sea *A* la conjunción de un número finito de generalizaciones empíricas acerca del aprendizaje formulado en términos de un vocabulario observacional V_B y sea *P* la conjunción de un número finito de oraciones arbitrarias formadas a partir de un conjunto V_T de términos no interpretados elegidos arbitrariamente (por ejemplo, podría ser la conjunción de los postulados de alguna axiomatización de la geometría elíptica). Entonces, tomando a *P* como postulados de *T* y eligiendo la oración $P \supset A$ como único miembro de nuestro sistema interpretativo *J*, obtenemos una teoría interpretada *T'*, que explica de un modo trivial todas las generalizaciones empíricas dadas, pues es evidente que $T \cdot J$ implica *A*. Con todo, es innecesario decir que *T'* no sería considerada satisfactoria como teoría del aprendizaje.[64] La característica que aquí se ilustra no invalida nuestro análisis de la interpretación parcial puesto que esta última no pretende que todo sistema teórico parcialmente interpretado sea en potencia una teoría interesante y, a decir verdad, incluso el requerimiento de definición completa de todos los términos teóricos por medio de V_B deja lugar a "teorías" igualmente insatisfactorias. Los ejemplos

[64] Es interesante advertir aquí que si una teoría interpretada tuviera que satisfacer además de las condiciones que se especificaron en la sección 8, los criterios de significación para los términos y las oraciones teóricas que han sido propuestos poco ha por Carnap (1956, secciones 6, 7, 8), los términos y las oraciones de nuestra ficticia "teoría del aprendizaje" serían desechados como no significativos.

similares al de nuestra ficticia "teoría del aprendizaje" sólo nos recuerdan que una buena teoría científica debe cumplir el requisito de tener una interpretación empírica (que es necesaria si la teoría ha de permitir consecuencias que puedan ponerse a prueba en forma empírica), y además debe satisfacer otras importantes condiciones: sus consecuencias V_B deben estar bien confirmadas en forma empírica, deben realizar una sistematización lógicamente simple de las oraciones V_B pertinentes, deben sugerir nuevas leyes empíricas, etcétera.

Si se considera a las oraciones de una teoría parcialmente interpretada T' como enunciados significativos, puede decirse que son verdaderas o falsas. Es posible, pues, enfrentar de lleno la cuestión de la referencia fáctica de los términos teóricos a la que aludimos antes en esta sección: afirmar que los términos de una teoría dada tienen referencia fáctica, que las entidades a las que pretenden aplicarse existen en la realidad, equivale a afirmar que lo que nos dice la teoría es verdadero y esto a su vez equivale a afirmar la teoría. Cuando decimos, por ejemplo, que las partículas elementales de la física teórica contemporánea existen realmente, estamos afirmando que en el universo existen partículas de los diversos tipos que indica la teoría física, que dichas partículas están regidas por determinadas leyes físicas y que señalan su presencia en determinadas circunstancias mediante ciertos tipos específicos de síntomas observables. Pero esto equivale a afirmar la verdad de la teoría física (interpretada) de las partículas elementales. De modo similar, afirmar la existencia de impulsos, reservas, fuerzas de hábito, etc., que postula una teoría del aprendizaje, determinada, lleva a afirmar la verdad del sistema que consta de los enunciados de la teoría y de su interpretación empírica.[65]

Entendida así, la existencia de las entidades hipotéticas con determinadas características e interrelaciones, tal como las supone una teoría dada, puede examinarse de manera inductiva, así como puede examinarse de manera inductiva la verdad de la teoría misma, mediante pruebas empíricas de sus consecuencias en términos de V_B.

De acuerdo con la concepción que se ha esbozado antes, tenemos que atribuir referencia fáctica a todos los términos (extralógicos) de una teoría, si es que esa teoría es verdadera; en consecuencia, esta característica no proporciona una base para una dicotomía semántica en el vocabulario teórico. Además, la referencia fáctica de los términos teóricos, tal como se la concibe aquí, no depende de que se los pueda evitar o no, sustituyéndolos por expresiones que constan de términos de V_B solamente.

[65] Más precisamente la oración de Ramsey asociada con T' expresa la afirmación de que existen entidades de diversos tipos (tales como objetos y acontecimientos hipotéticos y sus diversas propiedades y relaciones cualitativas y cuantitativas) postuladas por una teoría interpretada T'. Se la obtiene reemplazando a todas las constantes teóricas en la conjunción de los postulados de T' por variables y ligando éstas mediante cuantificadores existenciales colocados al comienzo de la expresión resultante. La oración así obtenida es una consecuencia lógica de los postulados de T'; pero la recíproca no vale. En consecuencia, estrictamente hablando, la afirmación de la existencia de las diversas entidades hipotéticas supuestas en una teoría es más débil, desde el punto de vista lógico, que la teoría misma.

Para observaciones interesantes sobre la cuestión de la realidad de las entidades teóricas, véase por ejemplo, Toulmin (1953, págs. 134-139) y Smart (1956).

Aun cuando se defina explícitamente en términos de V_B a todos los términos teóricos de una teoría T', de manera tal que su uso proporcione un modo abreviado conveniente de decir lo que también podría decirse por medio de V_B solamente, los términos teóricos tendrán aún referencia fáctica si lo que dice la teoría es verdadero.

Las observaciones precedentes acerca de las verdad y la referencia fáctica con respecto a las teorías parcialmente interpretadas suponen que se ha considerado a las oraciones como enunciados significativos. Para los que encuentran que esta suposición es inaceptable, hay por lo menos otros dos modos de concebir lo que hemos llamado una teoría interpretada. El primero es el método de Ramsey, descripto en la sección anterior, que tiene la atrayente característica de representar a una teoría interpretada como un enunciado de buena fe, que no contiene otras constantes extralógicas excepto las que pertenecen a V_B: y que tienen exactamente las mismas consecuencias V_B que la teoría enunciada en términos de constantes teóricas interpretadas en forma incompleta. Es quizás el modo más satisfactorio de concebir el carácter lógico de una teoría científica. Pero la encontrarán objetable en especial, o quizá sólo ellos, los que por motivos filosóficos se oponen a ciertos compromisos ontológicos.[66] Esos compromisos son los que acarrea la aprobación de todas las variables requeridas cuyos dominios de aplicación son diferentes del que constituyen los individuos de la teoría (tales como, por ejemplo, el conjunto de todas las características cuantitativas de los objetos físicos o el conjunto de todas las relaciones diádicas entre ellos, o los conjuntos de esos conjuntos y así sucesivamente).

Finalmente, los que como los nominalistas contemporáneos rechazan esos compromisos ontológicos, podrán concebir a las teorías científicas, no como enunciados significativos sino como intrincados artificios para inferir, a partir de enunciados iniciales inteligibles. expresados en términos de un vocabulario V_B, previamente comprendido, otros enunciados también inteligibles, en términos de ese mismo vocabulario.[67] Así, los que se inclinan hacia el nominalismo podrán concebir los términos teóricos como signos auxiliares sin sentido que servirán como artificios simbólicos convenientes en la transición de un conjuto de enunciados empíricos a otro. Con seguridad, la concepción de leyes y teorías como principios extralógicos de inferencia no refleja el modo en que las usan los científicos teóricos.

[66] Se usa aquí el concepto en el sentido de Quine. De acuerdo con éste, una teoría tiene un compromiso ontológico con las entidades que deben incluirse en los dominios sobre los cuales se sitúan las variables ligadas, si es que la teoría ha de ser verdadera. Quine desarrolla y define esta idea en varios de los ensayos que comprenden su libro (1953).

[67] Varios autores han sugerido esta concepción de las leyes o las teorías como principios de inferencia pero no, en general, desde un punto de vista nominalista. Entre ellos puede citarse a Schlick (1931, págs. 151 y 155), Ramsey (1931), pág. 241; Ryle (1949, en especial, págs. 120-125); y Toulmin (1953) caps. III y IV. Toulmin hace notar sin embargo, que pensar que las leyes de la naturaleza son reglas o licencias "refleja sólo una parte de su naturaleza" (*loc. cit.*, pág. 105). Véase también la exposición de Braithwaite, sobre el tema (1953), págs. 85-87. Finalmente, el ensayo de Popper (1956) contiene varios comentarios críticos y constructivos que versan sobre ésta y otras cuestiones expuestas en el presente ensayo.

En las publicaciones que se ocupan de problemas de física teórica, o de biología o de psicología, por ejemplo, es normal tratar a las oraciones que contienen términos teóricos igual que a las que sirven para describir datos empíricos: funcionan junto con éstos como premisas y conclusiones de argumentos deductivos e inductivos. A decir verdad, para el científico en ejercicio, el uso y formulación real de los principios teóricos como complejas reglas extralógicas de inferencia sería un obstáculo más que una ayuda. Sin embargo, el propósito de los que sugieren esta concepción no es por supuesto facilitar el trabajo del científico sino más bien esclarecer el alcance de sus formulaciones; desde el punto de vista de un analista filosófico con inclinaciones nominalistas el enfoque propuesto de las oraciones científicas que, de acuerdo con sus pautas no son admisibles como enunciados, representa un avance en el esclarecimiento.

Sin embargo, la cuestión planteada por el dilema del teórico puede suscitarse también con respecto a las dos concepciones diferentes de las teorías. En cuanto a la formulación de Ramsey, cabe preguntar si no es posible prescindir del todo de las variables con cuantificadores existenciales que representan a los términos teóricos, evitando así el compromiso ontológico que requieren, sin sacrificar ninguna de las conexiones deductivas entre las oraciones V_B que establece la oración de Ramsey. Con respecto a las teorías concebidas como artículos para hacer inferencia, cabe preguntar si no se las puede reemplazar por un conjunto funcionalmente equivalente de reglas —o sea, un conjunto que establezca exactamente las mismas inferencias entre las oraciones V_B— que no use ningún "signo sin sentido".

En ambos casos, el teorema de Craig da una respuesta afirmativa, proporcionando un método general para construir el tipo deseado de equivalente. Pero en ambos casos, el resultado tiene los mismos defectos que se mencionaron en la sección 8. Primero, el método reemplazaría a la oración de Ramsey por un conjunto infinito de reglas, en términos de V_B y conduciría así a una pérdida de economía. Segundo, el sistema resultante de postulados o de reglas de inferencia no se prestaría para la explicación y la predicción inductiva. Tercero, tendría el defecto pragmático, ya puesto de manifiesto parcialmente en el segundo punto, de ser menos fructífero heurísticamente que el sistema que usa términos teóricos.

Nuestro argumento (5.1), el dilema del teórico, consideró que el único propósito de una teoría era establecer conexiones deductivas entre oraciones observacionales: si éste fuera el caso, los términos teóricos serían entonces innecesarios. Pero si se reconoce que una teoría satisfactoria debe proporcionar también posibilidades para las predicciones inductivas y que debe lograr una economía sistemática, así como una fertilidad heurística, resulta claro que no se pueden reemplazar las formulaciones teóricas por expresiones en términos de observables solamente; como se ha visto, el dilema del teórico, cuya confusión afirma lo contrario, parte de una premisa falsa.

Bergmann, Gustav: "The Logic of Psychological Concepts", *Philosophy of Science,* 18:93-110 (1951).

Bergmann, Gustav: "Comments on Professor Hempel's 'The Concept of Cognitive Significance'", *Proceedings of the American Academy of Arts and Sciences,* 80 (Nº 1): 78-86 (1951). Reimpreso en Gustav Bergmann, *The Metaphysics of Logical Positivism.* Nueva York, Longmans, Green and Co., 1954. Se hace referencia a este ensayo como (1951a).

Bergmann, Gustav y Kenneth Spence: "Operationism and Theory in Psychology", *Psychological Review,* 48:1-14 (1941). Reimpreso en Marx (1951).

Braithwaite, R. B.: *Scientific Explanation.* Cambridge, Inglaterra, Cambridge University Press, 1953.

Bridgman, P. W.: *The Logic of Modern Physics.* Nueva York, Macmillan, 1927.

Burks, Arthur W.: "The Logic of Causal Proposition", *Mind,* 60:363-382 (1951).

Campbell, Norman R.: *Physics: The Elements.* Cambridge, Cambridge University Press, 1920. Publicado nuevamente con el título *Foundations of Science.* Nueva York, Dover, 1957.

Carnap, Rudolf: "Testability and Meaning", *Philosophy of Science,* 3:420-468 (1936); 4:1-40 (1937). Reimpreso como monografía por Whitlock's Inc., New Haven, Conn., 1950. Extractos reeditados en Feigl and Brodbeck (1953).

Carnap, Rudolf: *The Logical Syntax of Language.* Londres, Routledge and Kegan Paul, 1937.

Carnap, Rudolf: *Foundations of Logic and Mathematics.* Chicago, University of Chicago Press, 1939.

Carnap, Rudolf: *Logical Foundations of Probability.* Chicago, University of Chicago Press, 1950.

Carnap, Rudolf: "Meaning Postulates", *Philosophical Studies,* 3:65-73 (1952).

Carnap, Rudolf: "Meaning and Synonymy in Natural Languages", *Philosophical Studies,* 6:33-47 (1955).

Carnap, Rudolf: "The Methodological Character of Theoretical Concepts", en H. Feigl y M. Scriven (comps.): *The Foundations of Science and the Concepts of Psychology and Psychoanalysis,* págs. 38-76. Minneapolis, University of Minnesota Press, 1956.

Cohen, M. R.y E. Nagel: *Introduction to Logic and Scientific Method,* Nueva York, Harcourt Brace, 1934.

Craig, William: "On Axiomatizability within a System", *Journal of Symbolic Logic,* 18:30-32 (1953).

Craig, William: "Replacement of Auxiliary Expressions", *Philosophical Review,* 63:38-55 (1956).

Cramér, Harald: *Mathematical Methods of Statistics.* Princeton, Princeton University Press, 1946.

Estes, W. K., S. Koch, K. MacCorquodale, P. E. Meehl, C. G. Mueller, W. S. Schoenfeld y W. S. Verplanck: *Modern Learning Theory.* Nueva York, Appleton-Century-Crofts, 1954.

Feigl, Herbert: "Some Remarks on the Meaning of Scientific Explanation". [Una versión levemente modificada de los comentarios publicada primero en *Psychological Review,* 52 (1948)], en Feigl y Sellars (1949), págs. 510-514.

Feigl, Herbert: "Existential Hypotheses", *Philosophy of Science,* 17:35-62 (1950).

Feigl, Herbert: "Logical Reconstruction, Realism, and Pure Semiotic", *Philosophy of Science,* 17:186-195 (1950). Mencionado en este ensayo como (1950s).

Feigl, Herbert: "Principles and Problems of Theory Construction in Psychology", en W. Dennis (comp.): *Current Trends in Psychological Theory,* págs. 179-213. Pittsburgh, University of Pittsburgh Press, 1951.

Feigl, Herbert y Wilfred Sellars (comps.): *Readings in Philosophical Analysis.* Nueva York, Appleton-Centry-Crofts, 1953.

Feigl, Herbert y Wilfred Sellars (comps.): *Readings in Philosophical Analysis.* Nueva York, Appleton-Century-Crofts, 1949.

Goodman, Nelson: "On Likeness of Meaning", *Analysis,* 10: 1-7 (1949). Reimpreso en una versión revisada en Linsky (1952).

Goodman, Nelson: *The Structure of Appearance.* Cambridge, Mass., Harvard University Press, 1951.
Goodman, Nelson: *Fact, Fiction, and Forecast.* Cambridge, Mass., Harvard University Press, 1955.
Grice, H. P. y P. F. Strawson: "In Defense of a Dogma", *Philosophical Review,* 65:141-158 (1956).
Hempel, Carl G.: "Geometry and Empirical Science", *American Mathematical Monthly,* 52-7-17 (1945). Reimpreso en Feigl and Sellars (1949) en Wiener (1953), y en James R. Newman (comp.): *The World of Mathematics.* Nueva York, Simon and Schuster, 1956.
Hempel, Carl G.: "A Note on Semantic Realism", *Philosophy of Science,* 17:169-173 (1950).
Hempel, Carl C.: "General System Theory and the Unity of Science", *Human Biology,* 23:313-322 (1951).
Hempel, Carl. G.: *Fundamentals of Concept Formation in Empirical Science,* Chicago, University of Chicago Press, 1952.
Hempel, Carl G.: "Implications of Carnap's Work for the Philosophy of Science", en F. A. Schilpp (comp.): *The Philosophy of Rudolf Carnap.* La Salle, Ill., Open Court Publishing Co., 1963.
Hempel, Carl G.: "The Concept of Cognitive Significance: A Reconsideration". *Proceedings of the American Academy of Arts and Sciences,* 80 (Nº 1):61-77 (1951).
Hempel, Carl G. y Paul Oppenheim: "Studies in the Logic of Explanation", *Philosophy of Science,* 15:135-175 (1948).
Hermes, H.: "Eine Axiomatisierung der allgemeinen Mechanik", *Forschungen zur Logik und Grundlegung der exakten Wissenschaften.* Neue Folge, Nº 3. Leipzig, 1938.
Hull, C. L.: "The Problem of Intervening Variables in Moral Behavior Theory", *Psychological Review,* 50:273-291 (1951). Reimpreso en Marx (1951).
Hull, C. L., C. I. Hovland, R. T. Ross, M. Hall, D. T. Perkins y F. B. Fitch: *Mathematic-Deductive Theory of Rote Learning.* New Haven, Yale University Press, 1940.
Hutten, Ernest H.: *The Language of Modern Physics. An Introduction to the Philosophy of Science.* Londres y Nueva York, Macmillan, 1956.
Kemeny, John G.: "Review of Carnap (1950)". *The Jornal of Symbolic Logic,* 16: 205-207 (1951).
Kemeny, John G.: "Extension of the Methods of Inductive Logic", *Philosophical Studies,* 3:38-42 (1952).
Kemeny, John G. y Paul Oppenheim: "On Reduction", *Philosophical Studies,* 7:6-19 (1956).
Lindzey, Gardner: "Hypothetical Constructs, Conventional Constructs, and the Use of Physiological Data in Psychological Theory", *Psychiatry,* 16:27-33 (1953).
Linsky, Leonard (comp.): *Semantics and the Philosophy of Language.* Urbana, Ill., University of Illinois Press, 1952.
MacCorquodale, K. y P. Meehl: "On a Distinction between Hypothetical Constructs and Intervening Variables", *Psychological Review,* 55:95-107 (1948). Reimpreso en Feigl y Brodbeck (1953) y con omisiones en Marx (1951).
MacCorquodale, K. y P. Meehl: "Edward C. Tolman", en Estes y col. (1954), págs. 177-266.
Margenau, Henry: *The Nature of Physical Reality.* Nueva York, McGraw-Hill Book Co., 1950.
Martin, R. M.: "On 'Analytic' ". *Philosophical Studies,* 3:42-47 (1952).
Marx, Melvin H. (comp.): *Psychological Theory.* Nueva York, Macmillan, 1951.
Mates, Benson: "Analytic Sentences", *Philosophical Review,* 60:525-534 (1951).
McKinsey, J. C. C., A. C.Sugar y P. Suppes: "Axiomatic Foundations of Classical Particle Mechanics", *Journal of Rational Mechanics and Analysis,* 2:253-272 (1953).
McKinsey, J. C. C. y P. Suppes: "Transformations of Systems of Classical Particle Mechanics", *Journal of Rational Mechanics and Analysis,* 2:273-289 (1953).
Mises, R. von: *Positivism: A Study in Human Understanding.* Cambridge, Mass., Harvard University Pres, 1951.

Nagel, Ernest: *Principles of the Theory of Probability*. Chicago, University of Chicago Press, 1939.
Nagel, Ernest: "The Meaning of Reduction in the Natural Sciences", en Robert C. Stauffer (comp.): *Science and Civilization*. Madison, Wis., University of Wisconsin Press, 1949. Reimpreso en Wiener (1953).
Neumann, John von y Oskar Morgenstern: *Theory of Games and Economic Behavior*. 2ª ed. Princeton, Princeton University Press, 1947.
Northrop, F. S. C.: *The Logic of the Sciences and the Humanities*. Nueva York, Macmillan, 1947.
Pap, Arthur: "Reduction Sentences and Open Concepts", *Methodos*, 5:3-28 (1953).
Pap, Arthur: *Analytische Erkenntnistheorie*, Viena, J. Springer, 1955.
Popper, Karl: *Logik der Forschung*. Viena, J. Springer, 1935.
Popper, Karl: *The Open Society and its Enemies*. Londres, G. Routledge & Sons, 1945.
Popper, Karl: "Three Views Concerning Human Knowledge", en H. D. Lewis (comp.): *Contemporary British Philosophy: Personal Statements*. Nueva York, Macmillan. 1956.
Quine, W. V.: "Two Dogmas of Empiricism", *Philosophical Review*, 60:20-43 (1951). Reimpreso en Quine (1953).
Quine, W. V.: *From a Logical Point of View*. Cambridge, Mass., Harvard University Press, 1953.
Ramsey, Frank Plumpton: *The Foundation of Mathematics and other Logical Essays*. Londres, Kegan Paul. Nueva York. Harcourt Brace, 1931.
Reichenbach, Hans: *Axiomatik der relativistischen Raum-Zeit-Lehre*. Braunschweig. F. Vieweg & Sohn, 1924.
Reichenbach, Hans: *Philosophie der Raum-Zeit-Lehre*. Berlín, W. de Gruyter & Co., 1928.
Reichenbach, Hans: *Philosophic Foundations of Quantum Mechanics*. Berkeley y Los Angeles, University of California Press, 1944.
Reichenbach, Hans: *The Theory of Probability*. Berkeley, University of California Press, 1949.
Reichenbach, Hans: *The Rise of Scientific Philosophy*. Berkeley y Los Angeles, University of California Press, 1951.
Reichenbach, Hans: *Nomological Statements and Admissible Operations*. Amsterdam. North Holland Publishing Co., 1954.
Rozeboom, William W.: "Mediation Variables in Scientific Theory". *Psychological Review*, 63:249-264 (1956).
Rubin. H. y P. Suppes: *Transformations of Systems of Relativistic Particle Mechanics*. Technical Report nº 2. Preparado para la Oficina de Investigación Naval. Stanford University, Stanford, 1953.
Russell, Bertrand: *Introduction to Mathematical Philosophy*. Londres y Nueva York. Macmillan, 1919.
Russell, Bertrand: *Mysticism and Logic*. Nueva York, W. W. Norton & Co., 1929.
Ryle, Gilbert: *The Concept of Mind*. Londres, Hutchinson's University Library, 1949.
Schlick, M.: "Die Kausalitaet in der gegenwaertigen Physik", *Die Naturwissenschaften*. 19:145-162 (1931).
Skinner, B. F.: *Science and Human Behavior*. Nueva York, Macmillan. 1953.
Smart, J. J. C.: "The Reality of Theoretical Entities", *Australasian Journal of Philosophy*, 34:1-12 (1956).
Spence, Kenneth W.: "The Nature of Theory Construction in Contemporary Psychology". *Psychological Review*, 51:47-68 (1944). Reimpreso en Marx (1951).
Tarski, Alfred: "Einige Methodologische Untersuchungen über die Definierbarkeit der Begriffe", *Erkenntnis*, 5:80-100 (1935). Traducción inglesa en Tarski (1956).
Tarski, Alfred: *Introduction to Logic and to the Methodology of Deductive Sciences*. Nueva York, Oxford University Press, 1941.
Tarski, Alfred: "The Semantic Conception of Truth". *Philosophy and Phenomenological Research*, 4:341-375 (1944). Reimpreso en Feigl y Sellars (1949) y en Linsky (1952).
Tarski, Alfred: *Logic, Semantics, Metamathematics*. Trad. por J. H. Woodger. Oxford, The Clarendon Pres, 1956.
Tolman, E. C.: "Operational Behaviorism and Current Trends in Psychology", *Pro-*

ceeding of the 25th Anniversary Celebration of the Inauguration of Graduate Study, Los Angeles. 1936. págs. 89-103. Reimpreso en Marx (1951).
Tolman, E. C., B. F. Ritchie y D. Kalish: "Studies in Spatial Learning. I. Orientation and the Short-Cut." *Journal of Experimental Psychology*, 36:13-24 (1946).
Toulmin, Stephen: *The Philosophy of Science*. Londres, Hutchinson's University Library, 1953.
Verplanck, W. S.: "Burrhus F. Skinner", en Estes y col. (1954), 267-316.
Wald, A: *On the Principles of Statistical Inference*. Notre Dame Press, 1942.
Walker, A. G.: "Foundations of Relativity: Parts I and II", *Proceedings of the Royal Society of Edinburgh*, 62:319-335 (1943-1949).
Walsh, W. H.: *An Introduction to Philosophy of History*. Londres, Hutchinson's University Library, 1951.
Wang, Hao: "Notes on the Analytic-Synthetic Distinction", *Theoria*, 21:158-178 (1955).
White, Morton G.: "The Analytic and the Sinthetic: An Untenable Dualism", en S. Hook (comp.): *John Dewey: Philosopher of Science and of Freedom*. Nueva York, Dial Press, 1950. Reimpreso en Linsky (1952).
White, Morton G.: *Toward Reunion in Philosophy*. Cambridge, Mass., Harvard University Press, 1956.
Wiener, Philip P. (comp.): *Readings in Philosophy of Science*. Nueva York, Scribner, 1953.
Woodger, J. H.: *The Axiomatic Method in Biology*. Cambridge, Cambridge University Press, 1937.
Woodger, J. H.: *The Technique of Theory Construction*. Chicago, University of Chicago Press, 1939.
Woodger, J. H.: *Biology and Language*. Cambridge, Cambridge University Press, 1952.

Hay traducciones castellanas totales o parciales de las siguientes obras:
Carnap, R.: "Significado y sinonimia en los lenguajes naturales", en Bunge, M. (comp.): *Antología semántica*, Buenos Aires, Nueva Visión, 1960.
Cohen, M. y Nagel, E.: *Introducción a la lógica y al método científico*, Buenos Aires, Amorrortu, 1968.
Goodman, N.: "Los condicionales contrafácticos" (cap. I de *Facts, Fiction, and Forecast*), en Bunge M. (comp.): *Antología semántica*, Buenos Aires, Nueva Visión, 1960.
Popper, K.: *La lógica de la investigación científica*, Madrid, Tecnos, 1972, 2ª ed.
Popper, K.: *La sociedad abierta y sus enemigos*, Buenos Aires, Paidós, 1957.
Reichenbach, H.: *La filosofía científica*, México, Fondo de Cultura Económica, 1953.
Russell, B.: *Misticismo y lógica*, Buenos Aires, Paidós, 1961.
Tarski, A.: *Introducción a la lógica y a la metodología de las ciencias deductivas*, Buenos Aires, Espasa Calpe, 1951.
Tarski, A.: "La concepción semántica de la verdad y los fundamentos de la semántica científica", en Bunge, M. (comp.): *Antología semántica*, Buenos Aires, Nueva Visión, 1960.

Cuarta Parte

LA EXPLICACION CIENTIFICA

Capítulo IX

La función de las leyes generales en la historia [1]

1.

Es una creencia bastante generalizada que la historia, a diferencia de las llamadas ciencias físicas, trata sobre la descripción de hechos particulares del pasado, antes que de la búsqueda de las leyes generales que regirían dichos sucesos. Como caracterización del tipo de problemas por los que algunos historiadores principalmente se interesan no puede negarse este criterio; por el contrario, como una formulación de la función teórica de las leyes generales en la investigación histórica científica, es sin duda inaceptable. Las consideraciones siguientes intentan fundamentar esta afirmación mostrando con algún detalle que las leyes generales tienen funciones totalmente análogas en la historia y en las ciencias naturales; que son un instrumento indispensable de la investigación histórica y que hasta constituyen la base común de diversos procedimientos considerados a menudo como propios de las ciencias sociales, a diferencia de las naturales.

Por ley general entendemos aquí un enunciado de forma condicional universal que puede confirmarse o rectificarse por hallazgos empíricos adecuados. El término "ley" sugiere la idea de que el enunciado en cuestión efectivamente ha sido confirmado por los elementos adecuados disponibles: como esta calificación en muchos casos es irrelevante para nuestros propósitos, emplearemos con frecuencia las expresiones "hipótesis de forma universal", o brevemente "hipótesis universal" en lugar de "ley general" y, si fuera necesario, estableceremos la condición de confirmación satisfactoria por separado. En este capítulo, una hipótesis universal se supone que afirma una regularidad del siguiente tipo: en todos los casos en donde un hecho de una clase específica *C* ocurre en un cierto lugar y tiempo, otro hecho de una clase específica *E* ocurrirá en un lugar y tiempo relacionados de un modo específico con el lugar y el tiempo de ocurrencia del primer suceso. (Se han elegido los símbolos "*C*" y "*E*" para sugerir los términos "causa" y "efecto", que a menudo, aunque no siempre, son aplicados a hechos relacionados por una ley del tipo señalado.)

[1] En el presente capítulo reproducimos, con la debida autorización, un artículo

2.

La función principal de las leyes generales en las ciencias naturales es conectar hechos en pautas a las que habitualmente se las denomina *explicación* y *predicción*.

2.1. La explicación de la ocurrencia de un hecho de una clase específica E en un cierto lugar y tiempo consiste, como generalmente se lo expresa, en indicar las causas o determinar los factores de E. La afirmación de que un conjunto de hechos, digamos de las clases C_1, C_2 ... C_n, ha originado el hecho que ha de explicarse, equivale a expresar que según ciertas leyes generales un conjunto de hechos de las clases mencionadas se acompañan regularmente de otro de la clase E. Así, la explicación científica del hecho en cuestión consiste en

1) un conjunto de enunciados que afirman la ocurrencia de ciertos hechos C_1, ... C_n en ciertos lugares y momentos;
2) un conjunto de hipótesis universales, según las cuales
   a) los enunciados de ambos grupos se encuentran razonablemente bien confirmados por pruebas empíricas;
   b) de ambos grupos de enunciados puede deducirse lógicamente la oración que afirma la ocurrencia del hecho E.

En una explicación física, el grupo 1) describiría las condiciones iniciales y extremas para la ocurrencia del hecho final; por lo general, diremos que el grupo 1) establece las *condiciones determinantes* para el hecho que se debe explicar, mientras que el grupo 2) contiene las leyes generales en que se basa la explicación, y que implican el enunciado de que siempre que ocurran hechos de la índole descripta en el primero grupo, tendrá lugar un hecho de la clase que debe explicarse.

> Ejemplo: sea el hecho de que debe explicarse el estallido del radiador de un automóvil durante una noche fría. Las oraciones del grupo 1) pueden establecer las siguientes condiciones iniciales y extremas: el vehículo permaneció en la calle durante toda la noche; su radiador, de hierro, se encontraba lleno de agua hasta el borde y con la tapa atornillada herméticamente. La temperatura, durante la noche, descendió a 4°C, al anochecer, a —4°C, en la mañana; la presión barométrica era normal; la presión que soporta el material del radiador es x. El grupo 2) contendría ciertas leyes empíricas tales como: a O°C, con presión atmosférica normal, el agua se congela; por debajo de 4°C, la presión de una masa de agua aumenta al descender la temperatura, si el volumen permanece constante o disminuye; cuando el agua se congela, la presión nuevamente aumenta. Por último,

---

publicado originariamente en *The Journal of Philosophy* (39, 1942, págs. 35-48) y en el cual se han introducido ligeras modificaciones.

este grupo tendría que incluir una ley cuantitativa referente al cambio de la presión del agua en función de su temperatura y volumen.

A partir de enunciados de estas dos clases puede deducirse por razonamiento lógico la conclusión de que el radiador estalló durante la noche: se ha establecido una explicación del hecho considerado.

2.2. Es importante recordar que los símbolos utilizados "*E*", "*C*", "C_1", y "C_2", etc., representan clases o propiedades de hechos y no lo que se suele denominar hechos individuales, porque el objeto de la descripción y explicación en todas las ramas de las ciencias empíricas es siempre la ocurrencia de un hecho de cierta *clase* (tal como un descenso de temparatura de 8°C, un eclipse lunar, una división celular, un terremoto, un aumento en el empleo, un asesinato político) en un cierto lugar y momento, o en un objeto empírico dado (tal como el radiador de un cierto automóvil, el sistema planetario, una personalidad histórica específica, etc.) en un determinado momento.

Lo que a veces se llama *descripción completa* de un hecho individual (tal como el terremoto de San Francisco en 1906 o el asesinato de Julio César) requeriría enunciar todas las propiedades exhibidas por la zona espacial o por el objeto individual implicado, durante el período ocupado por el hecho en cuestión. Esa tarea nunca puede cumplirse por completo.

A fortiori, es imposible dar una *explicación completa* de un hecho individual en el sentido de poder explicar *todas* sus características mediante hipótesis universales, aunque la explicación de lo acontecido en un lugar y momento específicos puede ser gradualmente más específica e inclusiva.

Pero, al respecto, no existe diferencia alguna entre la historia y las ciencias naturales: ambas explican sus temas sólo en términos generales, y la historia puede "captar la individualidad singular" de sus objetos de estudio ni más ni menos que la física o la química.

3.

Los puntos siguientes resultan más o menos directamente del estudio efectuado de la explicación científica y tienen especial importancia para los problemas que se analizarán.

3.1. Se puede decir que un conjunto de hechos ha provocado el hecho que debe explicarse, sólo si se pueden indicar las leyes generales que vinculan las "causas" y el "efecto" de la manera antes mencionada.

3.2. Al margen de que se utilice o no la terminología de causa-efecto, sólo se logra una explicación científica si se han aplicado leyes empíricas del tipo señalado en (2) o en 2.1.[2]

[2] Maurice Mandelbaum, en su análisis esclarecedor de la relevancia y causación en la historia (*The Problem of Historical Knowledge* [El problema del conocimiento histórico], Nueva York, 1938, caps. 7 y 8) parece sostener que hay una diferencia

3.3. El empleo de hipótesis empíricas universales como principios explicativos distingue la explicación genuina de la seudoexplicación tal como, digamos, el intento de explicar ciertos rasgos de la conducta orgánica haciendo referencia a una entelequia, respecto de cuyo funcionamiento no se ofrecen leyes; o la explicación de los logros de una determinada persona en función de su "misión histórica", su "sino predestinado" o nociones semejantes. Las explicaciones de este tipo se basan en metáforas más que en leyes; transmiten impresiones pictóricas y emocionales en lugar del insight en conexiones fácticas; sustituyen por analogías imprecisas y "plausibilidad" intuitiva las deducciones a partir de enunciados comprobables y, por consiguiente, son inaceptables como explicaciones científicas.

Toda explicación de carácter científico puede someterse a pruebas objetivas; éstas incluyen

a) una comprobación empírica de las oraciones que establecen condiciones determinantes;
b) una comprobación empírica de las hipótesis universales sobre las cuales se basa la explicación;
c) una investigación de si la explicación es lógicamente concluyente, en el sentido de que la oración que describe los hechos que han de explicarse se infiere de los enunciados de los grupos 1) y 2).

4.

La función de las leyes generales en la *predicción científica* puede establecerse ahora de manera breve. Con mucha amplitud, la predicción en las ciencias empíricas consiste en derivar un enunciado sobre un cierto hecho futuro (por ejemplo, la posición relativa de los planetas respecto del sol en una fecha futura) de: 1. enunciados que describen ciertas condiciones pasadas o presentes conocidas (por ejemplo, las posiciones y momentos de los planetas en un instante pasado o presente), y 2. leyes generales adecuadas (por ejemplo, las leyes de la mecánica celeste). Así, la estructura lógica de una predicción científica es la misma que la de una explicación científica, tal como se la ha descripto en 2.1. En particular, en toda ciencia empírica, tanto la predicción como la explicación implican la referencia a hipótesis empíricas universales.

La distinción corriente entre explicación y predicción reside principalmente en una diferencia pragmática entre ambas: mientras en el caso de la explicación se sabe que el hecho final ha acontecido y deben buscarse, por lo tanto, sus condiciones determinantes, la situación se invierte en el

---

entre el "análisis causal" o "explicación causal" de un hecho y el establecimiento de las leyes científicas que lo gobiernan en el sentido indicado en el contexto. Señala que "las leyes científicas sólo pueden formularse sobre la base de un análisis causal", pero "que no son sustitutos de explicaciones causales completas" (*op. cit.*, pág. 238). Por las razones antes indicadas, esta distinción no parece justa: toda "explicación causal" es una "explicación por leyes científicas"; porque de ninguna manera que no sea la de referirse a leyes empíricas se puede sostener científicamente la conexión causal entre determinados hechos.

caso de la predicción. Aquí están dadas las condiciones iniciales y deben determinarse sus "efectos", que en los casos típicos aún no se han producido.

En vista de la igualdad estructural de la explicación y de la predicción, puede decirse que una explicación, como la caracterizada en 2.1., es incompleta a menos que pueda funcionar también como predicción: si el hecho final puede derivarse de las condiciones iniciales y de las hipótesis universales establecidas en la explicación, entonces de igual modo se habrán predicho antes de que aconteciera en realidad, tomando como base el conocimiento de las condiciones iniciales y de las leyes generales. Así, por ejemplo, las condiciones iniciales y las leyes generales que el astrónomo aduciría para explicar un cierto eclipse de sol, son las mismas que le habrían servido como base adecuada para predecir el eclipse antes de que aconteciese.

No obstante, sólo raramente, si es que nunca, las explicaciones son enunciadas de manera tan completa como para exhibir su carácter predictivo (que la prueba mencionada en c en 3.3. serviría para revelar). Es muy común que la explicación de la ocurrencia de un hecho sea incompleta. Así, podemos escuchar la explicación de que el fuego destruyó un granero "porque" se arrojó un cigarrillo encendido en el heno, o que un determinado movimiento político tiene un éxito espectacular "porque" aprovechó los prejuicios raciales tan difundidos. Del mismo modo, en el caso del radiador roto, la manera corriente de formular la explicación se limitaría a señalar que el vehículo fue abandonado al frío intenso y que el radiador estaba lleno de agua. En enunciados explicativos como éstos, las leyes generales que confieren a las condiciones declaradas el carácter de "causas" o "factores determinantes" son omitidas por completo (a veces, quizá como "rutina") y, además, la enumeración de las condiciones determinantes del grupo 1) es incompleta. Esto se ilustra con los ejemplos anteriores y también por el análisis primero del caso del radiador roto: como lo revelaría un examen más detenido, incluso la formulación mucho más minuciosa de las condiciones determinantes y de las hipótesis universales requeriría una ampliación con el objeto de que pudieran utilizarse como bases adecuadas para decidir la conclusión de que el radiador estalló durante la noche.

5.

En ciertos casos, la imperfección de una explicación determinada puede considerarse no esencial. Así, por ejemplo, podemos pensar que la explicación para el último ejemplo podría completarse si así lo deseáramos, porque tenemos razones para suponer que conocemos la clase de condiciones determinantes y de leyes generales que son relevantes en este contexto.

Con mucha frecuencia, sin embargo, tropezamos con explicaciones cuya imperfección no puede simplemente desecharse como no esencial. Las consecuencias metodológicas de esta situación se analizarán más adelante (especialmente en los parágrafos 5.3 y 5.4).

5.1. Las consideraciones anteriores se aplican tanto a las *explicaciones históricas* como a cualquiera otra rama de la ciencia empírica. La explicación histórica, además, aspira a demostrar que el hecho en cuestión no fue "por azar", sino que podría esperarse en vista de ciertos antecedentes o condiciones simultáneas. La expectativa aludida no es una profecía o una adivinación, sino la anticipación científica racional fundada en la presunción de leyes generales.

Si este concepto es correcto, parecería extraño que mientras la mayoría de los historiadores sugieren, por cierto, explicaciones de hechos históricos, muchos de ellos niegan la posibilidad de recurrir a algunas leyes generales en la historia. Sin embargo, es posible explicar esta situación por medio de un estudio más detallado de la explicación en la historia, como se aclarará en el curso del análisis siguiente.

5.2. En algunos casos, las hipótesis universales que subyacen a una explicación histórica son expuestas más bien de modo explícito, como lo ilustran los trozos en bastardilla del párrafo que se transcribe para explicar la tendencia de organismos gubernamentales a perpetuarse y extenderse:

> A medida que las actividades del gobierno se extienden, mayor cantidad de personas desarrollan un interés profesional en la continuación y expansión de las funciones gubernamentales. *La gente que tiene cierto trabajo no quiere perderlo; los que se han acostumbrado a ciertas habilidades no reciben bien los cambios; quienes se han habituado a ejercer cierto tipo de poder no desean abandonar su control —más bien, quieren desarrollar mayor poder y, en consecuencia, mayor prestigio...*
>
> Así, las oficinas e instituciones gubernamentales, una vez creadas, no sólo intentan fortalecerse contra todo ataque, sino también ampliar el ámbito de sus operaciones.[3]

La mayor parte de las explicaciones ofrecidas en historia o en sociología, sin embargo, fracasan en incluir una enunciación explícita de las regularidades generales que se suponen; parecería haber, por lo menos, dos razones para explicar este hecho: *primero,* las hipótesis universales en cuestión se relacionan con frecuencia con la psicología individual o social, con la que de una u otra manera, se supone que están familiarizados todos a través de la experiencia cotidiana; por lo tanto, se dan tácitamente por sentadas. Es ésta una situación bastante similar a la caracterizada en la sección 4.

Segundo, muy a menudo es difícil formular los supuestos que subyacen explícitamente, con suficiente precisión y de manera simultánea, de modo que concuerden con todas las pruebas empíricas relevantes de que se dispone. Es sumamente instructivo, al examinar la adecuación de una explicación propuesta, intentar reconstruir las hipótesis universales sobre

[3] Donald W. McConnell y otros: *Economic Behavior.* Nueva York, 1939, págs. 894-895 (la bastardilla es mía).

las que se fundamenta. En especial, términos tales como "por lo tanto", "por consiguiente", "en consecuencia", "porque", "naturalmente", "obviamente", etc., a menudo indican el supuesto tácito de alguna ley general: se emplean para vincular las condiciones iniciales con el hecho que debe explicarse; pero que éste deba esperarse "naturalmente" como una "consecuencia" de las condiciones establecidas, solamente se sigue si se presuponen leyes generales adecuadas. Consideremos, por ejemplo, la enunciación de que los agricultores del Dust Bowl emigraron a California "porque" la sequía y las tormentas de arena continuas hacían cada vez más precaria su existencia, y porque California parecía brindarles mejores condiciones de vida. Esta explicación se basa en una hipótesis universal según la cual la población tendería a emigrar a regiones que ofrecen mejores condiciones para vivir. Pero es obvio que sería difícil exponer esta hipótesis con precisión bajo la forma de una ley general que esté razonablemente verificada por todos los datos relevantes disponibles. Análogamente, si se explica una revolución particular haciendo referencia al descontento creciente de gran parte de la población en ciertas y determinadas condiciones, está claro que en esta explicación se supone una regularidad general; difícilmente, sin embargo, podremos establecer con exactitud de qué extensión y qué forma específica debe revestir el descontento y cuáles han de ser las condiciones ambientales para provocar una revolución. Observaciones similares se aplican a todas las explicaciones históricas en función de la lucha de clases, de las condiciones económicas o geográficas, de los intereses de ciertos grupos, de las tendencias al consumo conspicuo, etc.: todas ellas se basan en el supuesto de hipótesis universales[4] que conectan ciertas características de la vida individual o grupal con otros; en muchos casos, el contenido de las hipótesis que se supone tácitamente en una explicación dada, puede sólo reconstruirse de manera aproximada.

5.3. Podría argüirse que los fenómenos cubiertos por el tipo de explicación que se acaba de mencionar son de carácter estadístico y que, por consiguiente, sólo se requiere suponer las hipótesis de probabilidad en sus explicaciones, de modo que el problema de las "leyes generales subyacentes" estaría basado en una premisa falsa. Y, por cierto, parece posible y justificable interpretar ciertas explicaciones históricas como fundadas en la presunción de hipótesis de probabilidad, más bien que en la de leyes generales "deterministas", es decir, leyes en forma de condicionales universales. Esta pretensión puede extenderse a muchas de las explicaciones ofrecidas en otros campos de la ciencia. Así, por ejemplo, si Tomasito cae enfermo de sarampión dos semanas después que su hermano y no ha estado en contacto con otras personas que padecían esa afección, aceptamos la explicación de que se contagió del hermanito. Ahora bien, hay una

[4] Lo que a veces se denomina erróneamente explicación mediante un cierto *concepto* es en realidad, en la ciencia empírica, una explicación en términos de *hipótesis universales* que contienen dicho concepto. Las "explicaciones" que implican conceptos que no funcionan en hipótesis empíricas comprobables —tales como la "entelequia" en la biología, "el destino histórico de una raza" o el autodesarrollo de la razón absoluta" en la historia— son simples metáforas sin ningún contenido cognitivo.

hipótesis general subyacente a esta explicación, pero es difícil que pueda decirse que constituye una ley general según la cual toda persona que no haya tenido sarampión se contagiará sin excepción si tiene contacto con alguien que sufra de dicha enfermedad; sólo puede afirmarse que se producirá el contagio con una alta probabilidad.

Muchas de las explicaciones brindadas en la historia parecen admitir un análisis de esta índole: si fueran formuladas plena y explícitamente establecerían ciertas condiciones iniciales y ciertas hipótesis de probabilidad,[5] de modo que la ocurrencia del hecho que ha de explicarse es sumamente probable si se producen las condiciones iniciales, en vista de esas hipótesis. Pero al margen de que las explicaciones históricas se interpretan como causales o probabilísticas, sigue siendo verdad que, en general, las condiciones iniciales, y especialmente las hipótesis universales implicadas, no se indican con claridad y no pueden completarse sin ambigüedades. (En el caso de las hipótesis de probabilidad, por ejemplo, los valores de probabilidad incluidos, en el mejor de los casos sólo se conocerán de manera muy aproximada.)

5.4. Lo que los análisis explicativos de hechos históricos ofrecen es, entonces, en la mayoría de los casos no una explicación en uno de los sentidos indicados, sino algo que puede llamarse un *esbozo de explicación*. Este consiste en una indicación más o menos vaga de las leyes y las condiciones iniciales consideradas relevantes, y necesita "completarse" con el fin de convertirse en una explicación hecha y derecha. Este completamiento requiere una investigación empírica más extensa, para la cual el esbozo sugiere la orientación. (Los esbozos explicativos son comunes también en campos ajenos a la historia; muchas explicaciones en psicoanálisis, por ejemplo, ilustran este punto.)

Es obvio que el esbozo explicativo no admite una comprobación empírica en la misma medida que la explicación completa; y, sin embargo, existe una diferencia entre un esbozo explicativo científicamente aceptable y una seudoexplicación (o un esbozo de seudoexplicación). Un esbozo explicativo científicamente aceptable necesita completarse con enunciados más específicos; pero apunta en la dirección en que se encontrarán esos enunciados, y la investigación concreta puede tender a confirmar o debilitar esas indicaciones. Es decir, puede demostrar que el tipo de condiciones iniciales sugeridas son realmente relevantes; o puede revelar que han de tomarse en cuenta factores de índole totalmente distinta con el fin de llegar a una explicación satisfactoria.

Este proceso de completamiento requerido por el esbozo explicativo, en general realizará el aumento gradual en la precisión de las formulaciones implicadas, pero en cada etapa de este proceso esas formulaciones

[5] E. Zilsel, en un interesante trabajo sobre "Physics and the Problem of Historico-Sociological Laws" [La física y el problema de las leyes histórico-sociológicas] (*Philosophy of Science*, 1941, vol. 8, págs. 576-579), sugiere que todas las leyes específicamente históricas son de naturaleza estadística similares a las "macroleyes" de la física. Las observaciones realizadas, sin embargo, no se limitan a leyes específicamente históricas, puesto que la explicación en la historia se apoya en gran medida en leyes no históricas (véase la sección 8 de este capítulo).

tendrán cierto contenido empírico: será posible indicar, al menos de manera aproximada, qué tipo de prueba sería importante para verificarlos y qué hallazgos tenderían a confirmarlos. En el caso de explicaciones o esbozos explicativos no empíricos, por otro lado —digamos, por referencia al destino histórico de cierta raza, o a un principio de justicia histórico— el empleo de términos empíricamente carentes de sentido hace imposible indicar, aun de modo aproximado, el tipo de investigación que tendrá relación con esas formulaciones y que podrían conducir a pruebas que confirmasen o debilitasen la explicación propuesta.

5.5. Al tratar de apreciar la seriedad de una explicación determinada, en primer término ha de intentarse reconstruir, de modo tan completo como sea posible, el argumento que constituye la explicación o el esbozo explicativo. En especial, es importante advertir cuáles son las hipótesis explicativas subyacentes y estimar su alcance y fundamento empírico. Resucitar los supuestos enterrados bajo las lápidas del "de ahí", "por tanto", "porque" y otros semejantes, a menudo revelará que la explicación ofrecida está pobremente fundamentada o es inaceptable en absoluto. En muchos casos, este procedimiento aclarará la falacia de pretender que se haya explicado un gran número de detalles de un hecho, cuando incluso aceptando una interpretación muy liberal, sólo se han explicado características muy vastas. Así, por ejemplo, las condiciones geográficas o económicas en las cuales vive un grupo pueden explicar ciertos rasgos generales o sus códigos artísticos o morales, pero aceptar esto no significa que los logros artísticos del grupo o su sistema moral han sido explicados en detalle. Esto implicaría que de una descripción sola de las condiciones geográficas o económicas dominantes puede deducirse, mediante leyes generales especificables, un relato detallado de ciertos aspectos de la vida cultural del grupo.

Un error vinculado consiste en entresacar uno entre varios grupos importantes de factores que serían expresados en las condiciones iniciales y pretender luego que el fenómeno en cuestión está "determinado" por este grupo exclusivamente, pudiendo explicarse entonces en sus mismos términos.

A veces, los adherentes a alguna escuela particular de explicación o interpretación histórica aducirán, como prueba en favor de sus enfoques, una predicción histórica adecuada realizada por un representante de su escuela. Pero, aunque el éxito predictivo de la teoría es ciertamente prueba relevante de su corrección, es importante asegurar que la predicción exacta sea de hecho obtenible mediante la teoría en cuestión. Sucede a veces que la predicción es en realidad una adivinanza ingenua influida por la concepción teórica de su autor, pero que no puede lograrse sólo mediante esa teoría. Así, el adherente de una "teoría" plenamente metafísica de la historia puede tener una fuerte inclinación hacia los desarrollos históricos y hacer predicciones correctas, que incluso formulará con la terminología de su teoría, aunque no pueden haberse logrado con ella. Precaverse contra tales casos seudoconfirmados sería una de las funciones del test (*c*) en 3.3.

6.

Hemos tratado de demostrar que en historia, tanto como en cualquier otra rama de la investigación empírica, la explicación científica sólo puede lograrse mediante hipótesis generales adecuadas, o por teorías que son cuerpos de hipótesis sistemáticamente relacionadas. Esta tesis es con claridad contraria a la opinión familiar de que la explicación genuina en historia se obtiene mediante un método que distingue típicamente a las ciencias sociales de las naturales, es decir, el método de la *comprensión empática*: el historiador, se dice, se imagina a sí mismo en el lugar de las personas implicadas en los hechos que desea explicar; trata de percibir, de la manera más completa posible, las circunstancias en las cuales actuaron y los motivos que influyeron sobre sus actos; y mediante esta autoidentificación imaginaria con sus héroes logra el conocimiento y, por ende, una explicación adecuada de los hechos sobre los que se interesa.

Este método de empatía es, sin duda, con frecuencia aplicado por legos y expertos en historia. Pero por sí solo no constituye una explicación. Más bien es esencialmente un recurso heurístico; su función reside en sugerir hipótesis psicológicas que puedan servir como principios explicativos en el caso considerado. En términos crudos, la idea subyacente a esta función es la siguiente: el historiador trata de descubrir cómo actuaría él mismo en ciertas condiciones y bajo motivaciones específicas de sus héroes; a título de ensayo generaliza sus hallazgos en una regla general que utiliza como principio explicativo para explicar las acciones de las personas implicadas. Este procedimiento a veces puede ser de ayuda desde el punto de vista heurístico, pero no garantiza la corrección de la explicación histórica a que conduce. Esta última depende más de la corrección fáctica de las generalizaciones que pudo haber sugerido el método de la comprensión.

7.

Tampoco es indispensable el empleo de este método para la explicación histórica. Un historiador, por ejemplo, puede ser incapaz de sentirse a sí mismo en el rol de una personalidad histórica paranoica, y sin embargo muy bien puede explicar sus actos haciendo referencia a los principios de la psicología patológica. Así, si el historiador puede o no identificarse con el héroe histórico, carece de importancia para la corrección de su explicación; lo que cuenta es la corrección de las hipótesis generales implicadas, sea que hayan sido sugeridas por la empatía o por un procedimiento estrictamente conductista. Gran parte del atractivo del "método de la comprensión" parece deberse al hecho de que tiende a presentar los fenómenos en cuestión como "plausibles" o "naturales" para nosotros;[6] esto se logra a menudo con metáforas persuasivas. Pero la clase de "comprensión" así lograda debe distinguirse claramente de la compren-

[6] Para una crítica de esta clase de plausibilidad véase Zilsel, *op. cit.*, págs. 577-578, y los caps. 7 y 8 en "Problems of Empiricism" [Problemas del empirismo], en

sión científica. En la historia, como en cualquier otra ciencia empírica, explicar un fenómeno consiste en subsumirlo bajo leyes generales empíricas; el criterio de su corrección no reside en que atraiga nuestra imaginación o que se presente en términos de analogías sugestivas, que deba parecer más o menos plausible —cosas que también pueden ocurrir en las seudoexplicaciones— sino exclusivamente en si reposa sobre supuestos empíricamente bien confirmados, relativos a sus condiciones iniciales y a leyes generales.

7.1. Hasta aquí hemos analizado la importancia de las leyes generales para la explicación y la predicción y para la llamada comprensión en la historia. Examinemos ahora de modo más breve otros procedimientos de investigación histórica que implican suponer hipótesis universales.

Intimamente relacionada con la explicación y la comprensión es la llamada *interpretación de fenómenos históricos* en función de algún enfoque o teoría particular. Las interpretaciones que en realidad se ofrecen en la historia consisten en subsumir los fenómenos en cuestión bajo una explicación o esbozo explicativo científico; o en un intento de subsumirlos bajo una idea general que no puede comprobarse de modo empírico. En el primer caso, la interpretación claramente es una explicación por medio de hipótesis universales; en el último caso, se reduce a una seudoexplicación que puede tener cierto atractivo emocional y evocar asociaciones pictóricas vívidas, pero que no fomenta nuestra comprensión de los fenómenos que se consideran.

7.2. Consideraciones análogas se aplican al procedimiento de indagar el "sentido" de ciertos hechos históricos; su alcance científico consiste en determinar qué otros hechos están relacionados de manera relevante con el que se estudia, sea como "causas" o como "efectos"; el enunciado de las conexiones relevantes asume nuevamente la forma de explicaciones o esbozos explicativos que implican hipótesis universales. Esto se observa con más claridad en la siguiente sección.

7.3. En la explicación histórica de ciertas instituciones sociales se acentúa grandemente el análisis del *desarrollo* de la institución hasta la etapa que se examina. Los críticos de este enfoque han objetado que una simple descripción de esta índole no es una explicación genuina. Este argumento puede otorgar un aspecto ligeramente diferente en función de las reflexiones anteriores: la descripción del desarrollo de una institución obviamente no es describir *todos* los hechos que en el tiempo lo precedieron; sólo se incluirán aquellos que sean "*relevantes*" para la formación de esa institución. Y que un hecho sea relevante para ese desarrollo no es cuestión de la actitud valorativa, sino es cuestión objetiva que depende de lo que a veces se ha llamado el análisis causal del surgimiento de la institución.[7] Ahora bien, el análisis causal de un hecho establece una explica-

la *International Encyclopedia of Unified Science*. Chicago, University Chicago Press, 1941, vol. II.

[7] Véase la detallada y clara exposición de este punto en el libro de M. Mandelbaum, caps. 6 y 8.

ción, y puesto que ello requiere hacer referencia a hipótesis generales, también lo requerirán los supuestos acerca de la relevancia y, por ende, el análisis adecuado del desarrollo histórico de una institución.

7.4. De manera similar el empleo de las nociones de *determinación* y *dependencia* en las ciencias empíricas, incluso la historia, implica la referencia a leyes generales.[8] Así, por ejemplo, podemos decir que la presión de un gas depende de su temperatura y volumen, o que la temperatura y el volumen determinan la presión, en virtud de la ley de Boyle. Pero a menos que no se establezcan explícitamente las leyes subyacentes, la afirmación de una relación de dependencia o de determinación entre ciertas magnitudes o características sólo implica, en el mejor de los casos, pretender que están vinculadas por alguna ley empírica inespecífica; y ello por cierto, es una afirmación muy pobre: si sólo sabemos que hay determinada ley empírica que relaciona dos magnitudes métricas (tales como la longitud y la temperatura de una barra de metal), ni siquiera podríamos estar seguros de que un cambio en una de ellas se acompañará de un cambio en la otra (porque la ley puede relacionar el mismo valor de la magnitud "dependiente" o "determinada" con diferentes valores de la otra) sino a lo sumo, de que al dar un cierto valor específico a una de las variables siempre se asociará el mismo valor en la otra. Y ello, por cierto, es mucho menos de lo que la mayoría de los autores quieren afirmar cuando hablan de determinación o dependencia en los análisis históricos.

Por lo tanto, la afirmación imprudente de que las condiciones económicas, geográficas o cualquiera otra "determinan" el desarrollo y cambio de los restantes aspectos de la sociedad humana, sólo tiene valor explicativo en la medida en que puede apoyarse en leyes explícitas que establezcan justamente qué tipo de cambio en la cultura humana acusará cambios específicos con regularidad, en las condiciones económicas, geográficas, etc. Sólo el establecimiento de leyes concretas puede completar la tesis general con un contenido científico, sujetarlo a comprobaciones empíricas y conferirle función explicativa. La elaboración de esas leyes con la mayor precisión posible parece ser claramente la dirección y comprensión científicas.

[8] De acuerdo con Mandelbaum, la historia, a diferencia de las ciencias físicas, no consiste "en formular leyes de las cuales el caso particular sea un ejemplo, sino en la descripción de los hechos en su relación real recíproca determinante, al observar los hechos como productos y productores de cambios" (*op. cit.*, págs. 13-14). Esta es, en efecto, una concepción cuya insostenibilidad ya fue señalada por Hume, es decir, que un examen cuidadoso de dos hechos específicos solos, sin referencia a casos similares y a regularidades generales, puede revelar que uno de ellos produce o determina el otro. Esta tesis no sólo contradice el sentido científico del concepto de determinación, que se basa claramente en el de la ley general, sino que incluso no provee criterios objetivos que puedan indicar la relación buscada de determinación o producción. Así, hablar de una determinación empírica, al margen de toda referencia a leyes generales, es recurrir a la metáfora sin contenido cognitivo.

8.

Las consideraciones desarrolladas en el presente capítulo son completamente neutrales respecto del problema de las *leyes específicamente históricas*: no suponen un modo concreto de distinguir las leyes históricas de las sociológicas u otras leyes, ni aceptan o niegan el supuesto de que puedan encontrarse leyes empíricas que, en algún sentido específico, sean históricas y estén confirmadas por la evidencia empírica.

Pero es digno mencionar aquí que las hipótesis universales a las cuales los historiadores se refieren, explícita o tácitamente, al ofrecer explicaciones, predicciones, interpretaciones, juicios de importancia, etc., son tomadas de *diversos* campos de la investigación científica, hasta tanto no sean generalizaciones precientíficas de experiencias diarias. Muchas de las hipótesis universales que subyacen a las explicaciones históricas, por ejemplo, por lo común serían clasificadas como leyes psicológicas, económicas, sociológicas y, en parte, quizás históricas; además, la investigación histórica con frecuencia debe recurrir a leyes generales establecidas en física, química y biología. Así, por ejemplo, la explicación de la derrota de un ejército haciendo referencia a la falta de alimentos, condiciones meteorológicas adversas, enfermedades, etc., se funda en el supuesto —por lo general, tácito— de esas leyes. Emplear anillos de crecimiento de los árboles para ubicar en el tiempo hechos históricos se basa en la aplicación de ciertas regularidades biológicas. Diversos métodos para comprobar la autenticidad de documentos, pinturas, monedas, etc., utilizan teorías físicas y químicas.

Los dos últimos ejemplos ilustran otro punto que es importante en este contexto: incluso si un historiador se propusiera limitar su investigación a una "descripción pura" del pasado, sin intentar ofrecer explicaciones o enunciados acerca de la relevancia y determinación, tendría de continuo que recurrir a leyes generales, puesto que el objeto de sus estudios sería el pasado —definitivamente inaccesible a su examen directo—. Tendría que establecer su conocimiento por métodos indirectos: recurriendo a hipótesis universales que conecten datos presentes con aquellos hechos pasados. Este hecho ha quedado parcialmente oscurecido porque algunas de las regularidades implicadas son tan familiares que no se consideran dignas de mención. Y también, por la costumbre de relegar las diversas hipótesis y teorías utilizadas para indagar acerca de los hechos pasados, a las "ciencias auxiliares" de la historia. Es muy probable que algunos historiadores que tienden a minimizar, si no negar, la importancia de las leyes generales en la historia, actúen impulsados por el sentimiento de que sólo "leyes genuinas históricas" tienen interés para la historia. Pero una vez que advierten que el descubrimiento de leyes históricas (en alguno de los sentidos específicos de esta noción tan vaga) no convertiría a la historia en metodológicamente autónoma e independiente de las otras ramas de la investigación científica, parecería que el problema de la existencia de las leyes históricas perdiese parte de su importancia.

Las observaciones señaladas en esta sección no son sino ejemplos especiales de dos principios más amplios de la teoría de la ciencia: en primer

término, la separación entre "descripción pura" y "generalización hipotética y construcción teórica" en las ciencias empíricas carece de fundamento; en la estructuración del conocimiento científico ambas están inseparablemente eslabonadas. Y segundo, es similarmente infundado y fútil tratar de delimitar con una línea precisa los diferentes campos de la investigación científica y el desarrollo autónomo de cada uno de ellos. La necesidad en la investigación histórica de emplear con amplitud las hipótesis universales de las cuales, por lo menos, la inmensa mayoría provienen de campos de investigación distinguidos por tradición de la historia, es sólo uno de los aspectos de lo que puede denominarse unidad metodológica de las ciencias empíricas.

CAPÍTULO X

La lógica de la explicación [1]

INTRODUCCION

Uno de los objetivos primordiales de toda ciencia empírica es explicar los fenómenos del mundo de nuestra experiencia y responder no sólo a los "¿qué?", sino también a los "¿por qué?" Si bien por un lado hay coincidencia general sobre este punto, por el otro existen considerables diferencias de opinión acerca de la función y las características esenciales de la explicación científica. Este capítulo intenta arrojar alguna luz sobre estos problemas, mediante una revisión elemental de la pauta básica de la explicación científica, y realiza un análisis subsiguiente más riguroso del concepto de ley y de la estructura lógica de los argumentos explicativos.

La parte I de este capítulo está dedicada a un examen elemental; la parte II contiene un análisis del concepto de emergencia; en la parte III se busca manifestar y esclarecer, de modo más riguroso, algunos de los problemas lógicos singulares y desconcertantes que se originan en el análisis elemental corriente de la explicación. La parte IV, finalmente, trata sobre la idea del poder explicativo de una teoría. A este respecto, desarrolla una definición explícita y presenta una teoría formal para el caso de un lenguaje científico de estructura lógica simple.

[1] Este trabajo fue publicado previamente en *Philosophy of Science*, vol 15, págs. 135-175 (Copyright © 1948 por The Williams and Wilkins Co., Baltimore 2, Md., E.U.A.). Lo reproducimos, ligeramente modificado, con la debida autorización.
Se originó en una serie de discusiones que sostuvo el doctor Hempel con el doctor Oppenheim. Luego lo publicaron en colaboración; las contribuciones individuales no pueden discriminarse en detalle. Sin embargo, la esencia de la parte IV y la formulación definitiva de todo el texto se deben al autor mencionado en primer lugar. Algunas ideas desarrolladas en la parte II fueron sugeridas por el amigo común de ambos autores, el doctor Kurt Grelling, en la profusa correspondencia mantenida. Grelling, junto con su mujer, fue víctima del terror nazi durante la Segunda Guerra Mundial. Al incluir en este texto por lo menos algunas de las contribuciones de Grelling, que se identifican de modo explícito, queremos hacer realidad su deseo de que sus conceptos e ideas sobre el tema no cayeran completamente en el olvido.
Los autores expresan su deuda de gratitud a los profesores Rudolf Carnap, Herbert Feigl, Nelson Goodman y W. V. Quine, por sus estimulantes discusiones y críticas constructivas.



1. *Algunos ejemplos*

Si sumergimos rápidamente un termómetro en agua caliente se produce una caída transitoria de la columna mercurial, seguida de una repentina elevación. ¿Cómo explicar este fenómeno? El aumento de temperatura, al principio afecta sólo al vidrio del termómetro que se dilata y proporciona un espacio mayor al mercurio, cuyo nivel baja en consecuencia. Pero tan pronto como el aumento de temperatura alcanza al mercurio, éste se dilata y como su coeficiente de dilatación es considerablemente mayor que el del vidrio, determina una elevación del nivel mercurial.

Esta explicación consta de dos tipos de enunciados. Los del primer tipo indican ciertas condiciones que se manifiestan antes del fenómeno que se ha de explicar; nos referiremos brevemente a ellas como condiciones antecedentes. En nuestro ejemplo, estas condiciones abarcan, entre otros, el hecho de que el termómetro sea un tubo de vidrio que contiene en su interior mercurio, y que se lo sumerge en agua caliente. Los enunciados del segundo tipo expresan ciertas leyes generales; en este caso están incluidas las leyes de la dilatación termal del mercurio y del vidrio, y un enunciado sobre la escasa conductividad termal del vidrio. Ambos conjuntos de enunciados, cuando se los formula de manera completa y adecuada, explican el fenómeno que se estudia: infieren la consecuencia de que el mercurio descenderá primero para elevarse después. Por consiguiente el hecho analizado se explica por las leyes generales, es decir, demostrando que se produjo de acuerdo con esas leyes y en virtud de la realización de ciertas condiciones antecedentes específicas.

Consideremos otros ejemplos. Desde un bote, la pala del remo sumergida se ve quebrada hacia arriba. El fenómeno se explica mediante leyes generales, en especial por la ley de refracción y la de que el agua es un medio ópticamente más denso que el aire —haciendo referencia a ciertas condiciones antecedentes, o sea que parte del remo está en el agua, parte en el aire y que el remo es prácticamente un trozo de madera recto. Así, la pregunta "¿*Por qué* sucede el fenómeno?" deberá interpretarse como "¿De acuerdo con qué leyes generales y condiciones antecedentes se produce el fenómeno?"

Hasta ahora hemos considerado sólo la explicación de hechos particulares que ocurren en cierto tiempo y en cierto lugar determinados. Pero la pregunta "¿Por qué?" puede formularse también con respecto a leyes generales. Así, en nuestro último ejemplo, podría preguntarse "¿Por qué la propagación de la luz se acomoda a la ley de refracción?" La física clásica responde en función de la teoría ondulatoria de la luz, es decir, afirmando que la propagación de la luz es un fenómeno de onda de cierto tipo general, y que todo fenómeno ondulatorio de ese tipo satisface la ley de refracción. De este modo, la explicación de una regularidad general con-

siste en subsumirla dentro de otra regularidad más inclusiva, o sea una ley más general.

De manera similar, la validez de la ley de Galileo sobre la caída de los cuerpos en la superficie terrestre, puede explicarse a partir de un conjunto más inclusivo de leyes, tales como las del movimiento y de la gravedad de Newton, además de otros enunciados acerca de hechos particulares, como la masa y el radio terrestres.

2. *Pauta básica de la explicación científica*

De los ejemplos anteriores podemos ahora extraer algunas características generales de la explicación científica. Dividimos la explicación en dos componentes principales: *explanandum* y *explanans*.[2] Por explanandum entendemos la oración que describe el fenómeno a explicar (y no el fenómeno mismo); el término explanans se refiere a la clase de aquellas oraciones que se aducen para dilucidar el fenómeno. Cómo señalamos antes, el explanans se encuentra en dos subclases; una contiene ciertas oraciones C_1, C_2, ... C_k que formulan condiciones antecedentes específicas; otra es un conjunto de oraciones L_1, L_2, ... L_r que representan leyes generales.

Para que una explicación sea sólida, sus componentes deben satisfacer ciertas condiciones de adecuación, que pueden dividirse en condiciones lógicas y empíricas. Para el análisis que sigue, será suficiente formular estos requisitos de modo ligeramente impreciso; en la parte II se presentará un análisis y un replanteo más riguroso y preciso de estos conceptos.

2.1. *Condiciones lógicas de la adecuación*

(R1) El explanandum debe ser una consecuencia lógica del explanans; dicho en otras palabras, el primero debe ser lógicamente deducible de la información contenida en el explanans, porque de lo contrario este último no podría constituir una base adecuada para el explanandum.

(R2) El explanans debe contener leyes generales exigidas realmente para la derivación del explanandum. Sin embargo, no consideramos necesario para una explicación firme que el explanans deba contener por lo menos un enunciado que no sea una ley, puesto que, para mencionar sólo una razón, seguramente desearíamos considerar como explicación la deducción de las regularidades generales que gobiernan el movimiento de las estrellas dobles a partir de las leyes de la mecánica celeste, aun cuando todos los enunciados del explanans sean leyes generales.

[2] Estos dos términos derivados del latín *explanare*, se prefirieron a los quizá más corrientes de "explicandum" y "explicans" con el objeto de reservar estos dos últimos para emplearlos en el contexto de la explicación de significado o análisis. Acerca de la *explicación* en este sentido, véase Carnap, 1953a, pág. 513.

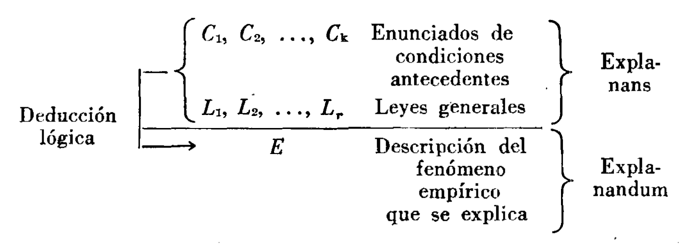
(R3) El explanans debe tener contenido empírico; es decir, que por lo menos en principio sea posible comprobarse mediante el experimento o la observación. Esta condición está implícita en (R1), pues, desde que se supone que el explanandum describe cierto fenómeno empírico, se puede concluir a partir de (R1) que el explanans entraña por lo menos una consecuencia de índole empírica, y este hecho le otorga la condición de ser verificable y de tener contenido empírico. Pero el punto merece una mención especial porque, como se verá en el parágrafo 3, ciertos argumentos que se ofrecen como explicaciones en las ciencias naturales y sociales violan esta exigencia.

2.2. *Condición empírica de la adecuación*

(R4) Las oraciones que constituyen el explanans han de ser verdaderas. Es obvio que en una explicación correcta los enunciados que constituyen el explanans deben satisfacer cierta condición de corrección fáctica. Pero parecería más adecuado estipular que el explanans ha de ser confirmado en alto grado por todos los elementos relevantes disponibles, antes que deba considerarse verdadero. No obstante, esta estipulación conduce a consecuencias embarazosas. Supóngase que en una etapa primitiva de la ciencia un determinado fenómeno fuera explicado mediante un explanans verificado con las pruebas de que se disponía en ese momento, pero que descubrimientos empíricos más recientes lo hubieran negado. En ese caso, deberíamos decir que originariamente la explicación fue correcta, pero que dejó de serlo cuando se descubrieron elementos de prueba desfavorables. Esto no parece concordar con el saludable uso común, que nos lleva a decir que basada en los elementos limitados de prueba iniciales, la verdad del explanans —y, por ende, la solidez de la explicación— había sido bastante probable, pero que la mayor evidencia ahora disponible hizo muy probable que el explanans no fuera verdadero; de ahí que la explicación no era ni había sido nunca, correcta.[3] (Expondremos e ilustraremos un punto similar con respecto al requisito de verdad para las leyes, al comenzar el parágrafo 5.)

Algunas de las características de la explicación que hasta ahora hemos indicado pueden resumirse en el siguiente esquema:

[3] (Agregado en 1964.) El requisito (R4) caracteriza lo que denominamos una *explicación verdadera* o correcta. En el análisis de la estructura lógica de los argumentos explicativos, por lo tanto, el requisito puede omitirse. Esto es, en efecto, lo que se ha hecho en sección 6 donde se introduce el concepto de ***explicación potencial.***



Se ha de señalar aquí que el mismo análisis formal, incluidas las cuatro condiciones necesarias, se aplica tanto a la predicción científica como a la explicación. La diferencia entre ambas es de carácter pragmático. Dado E, es decir, si sabemos que ha ocurrido el fenómeno descripto por E, y si se proporciona luego un conjunto adecuado de enunciados C_1, $C_2, \ldots, C_k, L_1, L_2, \ldots, L_r$, hablamos de una explicación del fenómeno que estudiamos. Si se proporcionan los últimos enunciados mencionados y se infiere E antes de que suceda el fenómeno que describe, hablamos de predicción. En consecuencia, puede decirse que la explicación de un hecho no es enteramente adecuada a menos que su explanans, considerado a tiempo, hubiera podido fundamentar el pronóstico del hecho que se analiza. Por lo tanto, lo expresado aquí sobre las características lógicas de la explicación, será aplicable a ambas, aunque sólo se mencione una de ellas.[4]

Sin embargo, por lo general —y en especial en el razonamiento precientífico— se brindan muchas explicaciones que carecen de esa fuerza potencial predictiva. Por ejemplo, puede decirse que un automóvil volcó en el camino "porque" uno de los neumáticos estalló cuando la velocidad del vehículo era alta. Está claro que con esta única información no hubiera podido predecirse el accidente, puesto que el explanans no suministra leyes generales explícitas por medio de las cuales se podría realizar la predicción, ni tampoco establece de manera adecuada las condiciones antecedentes que serían necesarias para ello. Idéntico punto puede ilustrarse haciendo referencia al criterio de W. S. Jevons de que toda explicación consiste en señalar una similitud entre hechos, y que a veces este proceso no necesita referirse a leyes en absoluto, y "tal vez no implique otra cosa que una única identidad, como cuando explicamos la aparición de estrellas fugaces señalando que son idénticas a las porciones de un cometa".[5] Pero es evidente que esa identidad no proporciona una explicación del fenómeno de las estrellas fugaces, a menos que presupongamos las leyes que gobiernan el desarrollo del calor y de la luz como efectos de la fricción. La observación de semejanzas posee un valor explicativo sólo si implica por lo menos alguna referencia tácita a las leyes generales.

En ciertos casos, los argumentos explicativos incompletos del tipo que ilustramos, simplemente suprimen partes del explanans por "obvias"; en otros, parecen implicar el supuesto de que mientras las partes que falten no sean obvias, el explanans incompleto podría, al menos, completarse con

[4] (Agregado en 1964.) Esta afirmación se examina en detalle, y se reafirma con ciertas calificaciones, en el capítulo IV.

[5] 1924, pág. 533.

el debido esfuerzo, como para posibilitar una estricta inferencia del explanandum. Esta suposición puede justificarse en ciertos casos, por ejemplo, cuando decimos que un terrón de azúcar desaparece "porque" lo sumergimos en una infusión caliente, pero no se podría justificar en muchos otros casos. Cuando se explican ciertos rasgos peculiares de la obra de un artista como desarrollos naturales de un tipo específico de neurosis, esta observación puede contener indicios significativos, pero en general no provee un fundamento suficiente para el pronóstico potencial de aquellos rasgos. En estos casos, la explicación incompleta puede considerarse, cuanto más, como un índice de cierta correlación positiva entre las condiciones antecedentes y el tipo de fenómeno que se ha de explicar, y como guía de la dirección que deberán tomar las investigaciones ulteriores, con el propósito de completar la explicación.

El tipo de explicación que hemos considerado hasta aquí comúnmente se denomina explicación causal.[6] Si E describe un hecho concreto, puede decirse entonces que las circunstancias antecedentes señaladas en las oraciones C_1, C_2, ... C_k "causan" en conjunto aquel hecho, en el sentido de que existen ciertas regularidades empíricas expresadas por las leyes L_1, L_2, ... L_r, las cuales implican que toda vez que ocurran condiciones del tipo indicado por C_1, C_2, ... C_k, tendrá lugar un hecho del tipo descripto en E. Los enunciados tales como L_1, L_2, ... L_r, que expresan conexiones generales y ordinarias entre características específicas de hechos, se denominan habitualmente leyes causales o deterministas. Estas leyes deberán distinguirse de las llamadas leyes estadísticas, las cuales expresan que, a la larga, un porcentaje explícitamente establecido de todos los casos que satisfagan un grupo dado de condiciones estará acompañado por un hecho de cierto tipo especificado. Ciertos casos de explicación científica implican la "subsunción" de un explanandum bajo un conjunto de leyes, de las cuales por lo menos algunas son del tipo estadístico. El análisis de la estructura lógica peculiar de esa clase de subsunción entraña problemas especiales y difíciles. El presente capítulo se limitará al examen del tipo deductivo de explicación, que todavía mantiene su significación en grandes sectores de la ciencia contemporánea, y aun en ciertas zonas donde una explicación más precisa exige referirse a leyes estadísticas.[7]

[6] (Agregado en 1964.) O bien, la explicación causal es una variedad del tipo deductivo.

[7] La descripción dada en el texto sobre las características generales de la explicación y de la predicción en la ciencia, no es de ningún modo novedosa; simplemente resume y enumera explícitamente algunos puntos fundamentales ya reconocidos por muchos científicos y metodólogos.

Así, por ejemplo, Mill afirma: "Se dice que un hecho individual se explica al señalarle su causa, vale decir, al formularse la ley o leyes causales de las cuales la producción del hecho es un ejemplo", y "se dice que una ley o uniformidad en la naturaleza, se explica cuando se señala otra ley o leyes de las cuales aquélla constituye sólo un caso, y de las cuales puede deducirse". (Libro III, cap. XII, sec. 1, 1858.) De modo similar, Jevons, cuya caracterización general de la explicación fue examinada críticamente en el texto, destaca que "el proceso más importante de explicación consiste en mostrar que un hecho observado es un caso de una ley o tendencia general" (1924, pág. 533). Ducasse afirma el mismo punto del modo siguiente: "La explicación consiste, esencialmente, en ofrecer una hipótesis acerca de un hecho, la

3. *La explicación en las ciencias no físicas. Enfoques teleológicos y motivacionales*

Nuestra caracterización de la explicación científica se ha basado hasta ahora en el estudio de casos tomados de las ciencias físicas. Pero los principios generales así obtenidos se aplican igualmente fuera de este campo.[8] Así, en psicología se explican varios tipos de conducta en animales de laboratorio y en seres humanos subsumiéndolos bajo leyes o aun teorías generales de aprendizaje o condicionamiento; y mientras con frecuencia las regularidades invocadas no pueden establecerse con la misma generalidad y precisión que en física o en química, al menos es evidente que el carácter general de aquellas explicaciones concuerda con nuestra primera caracterización.

Consideremos un ejemplo que incluya factores sociológicos y económicos. En el otoño de 1946 se produjo en los mercados algodoneros de los Estados Unidos una baja tan aguda del precio, que los mercados de Nueva York, Nueva Orleans y Chicago tuvieron que suspender sus actividades transitoriamente. En la tentativa de explicar este hecho, la prensa norteamericana lo investigó hasta descubrir que un especulador en gran escala de Nueva Orleans, temiendo que sus acopios fuesen excesivos, había comenzado a venderlos; los especuladores menores, presas del pánico, habían seguido su ejemplo, con lo cual precipitaron la crisis. Sin proponernos evaluar los méritos del argumento, señalamos que la explicación que aquí se sugiere, nuevamente implica enunciados sobre condiciones antecedentes y el supuesto de regularidades generales. Las primeras incluyen los hechos de que el primer especulador tenía mucha producción de algodón, que había especuladores menores con considerables acopios del material, que existían instituciones tales como los mercados algodoneros con su modo específico de operar, etc. Las regularidades generales referidas —según suele ocurrir en las explicaciones semipopulares— no se mencionan explícitamente, pero hay una implicación obvia de alguna forma de la ley de la oferta y la demanda que explica la baja de los precios del algodón en función del gran aumento de la oferta con una demanda prácticamente estable; además, es necesario contar con ciertas regularidades en la conducta de los individuos que tratan de mantener o mejorar su posición eco-

---

cual se encuentra, respecto de éste, como antecedente de un caso consecuente de alguna ley conectiva ya conocida" (1925, págs. 150-151).

Popper realizó un lúcido análisis de la estructura fundamental de la explicación y de la predicción (1935, sec. 12, y en una versión mejorada de 1945, especialmente en el capítulo 25 y en la nota 7 de ese mismo capítulo). Para una caracterización de la explicación como subsunción en teorías generales, véase por ejemplo el estudio sucinto de Hull (1943a, cap. 5). Un examen claro y elemental de ciertos aspectos de la explicación lo da Hospers (1946), y un examen conciso de muchos puntos esenciales en la explicación científica, que se consideran en las dos primeras partes del presente capítulo, se halla en Feigl (1945, págs. 284 y sigs.).

[8] Sobre el tema de la explicación en las ciencias sociales, especialmente en la historia, véanse también las siguientes publicaciones, que pueden servir para completar y ampliar el breve análisis que aquí se presenta: Hempel (1942), Popper (1945), White (1943); y los artículos *Cause* y *Understanding* en Beard y Hook (1946).

nómica. Por ahora tales leyes no pueden formularse con precisión y generalidad satisfactorias y, por ende, la explicación sugerida es con seguridad incompleta; pero su intención es sin lugar a dudas explicar el fenómeno integrándolo a una pauta general de regularidades económicas y sociopsicológicas.

Nos ocuparemos ahora de un argumento explicativo tomado del campo de la lingüística.[9] Al norte de Francia existe una variedad de sinónimos de la palabra "abeja", mientras que en la zona meridional hay esencialmente una sola palabra. Ante esta discrepancia, se ha sugerido la explicación de que durante la dominación romana, Francia meridional usaba la palabra "apícula", y los septentrionales, la palabra "apis". Esta última por un proceso de decadencia fonológica en el norte de Francia se transformó en el monosílabo "é"; los monosílabos tienden a eliminarse, en especial si contienen pocos elementos consonánticos, porque suelen provocar confusión. Así, se seleccionaron otros vocablos para evitarla. Pero "apícula", reducida a "abelho", permaneció suficientemente clara y se mantuvo para luego incorporarse al lenguaje común bajo la forma de "abeille". Si bien la explicación que se describe aquí es incompleta en el sentido caracterizado en la sección anterior, es evidente que hace referencia a condiciones antecedentes específicas tanto como a leyes generales.[10]

Mientras que los ejemplos de este tipo tienden a apoyar el criterio de que la explicación en biología, en psicología y en las ciencias sociales tiene la misma estructura que en las ciencias físicas, está bastante difundida la opinión de que, en muchos casos, el tipo causal de explicación es esencialmente inadecuado en otros campos que no sean la física y la química, y especialmente en el estudio de la conducta intencional. Examinemos con brevedad algunas de las razones que se han aducido en apoyo de este concepto.

Una de las más familiares entre ellas es la idea de que los hechos que implican actividades humanas, realizadas de manera individual o en grupo, tienen una singularidad peculiar y sin repetición que los hace inaccesibles a la explicación causal porque ésta, al confiar en uniformidades, presupone la repetibilidad de los fenómenos en consideración. Este argumento que incidentalmente también se utilizó para sostener la afirmación de que el método experimental es inaplicable en psicología y en las ciencias sociales, implica una falta de comprensión del carácter lógico de la explicación causal. Todo hecho individual es único, sea en las ciencias físicas, en psicología o en ciencias sociales, en el sentido de que no se repite con

[9] El ejemplo está tomado de Bonfante (1946, sec. 3).

[10] Mientras en cada uno de los dos últimos ejemplos, incuestionablemente ciertas regularidades confían en el argumento explicativo, no es posible argumentar de modo convincente que esas leyes a que se ha querido hacer referencia y que hoy no pueden formularse explícitamente, tengan un carácter causal antes que estadístico. Es bien posible que la mayoría o todas las regularidades que serán descubiertas con el desarrollo de la sociología, sean de carácter estadístico. Véanse sobre este tema las sugestivas observaciones de Zilsel (1941, sec. 8 y 1941a). Este problema no afecta, sin embargo, el punto principal que queremos señalar aquí, es decir, que en las ciencias sociales no menos que en las físicas, la subsunción en regularidades generales es indispensable para la explicación y la comprensión teórica de todo fenómeno.

todas sus características peculiares. Sin embargo, los hechos individuales pueden acomodarse a leyes generales del tipo causal y ser explicados por ellos, porque todo lo que afirma la ley causal es que todo hecho de índole específica, es decir, que reúna ciertas características determinadas se acompaña de otro que, a su vez, tiene ciertos rasgos específicos. Por ejemplo, todo hecho que implique fricción genera calor. Y todo lo que se requiere para que esas leyes puedan comprobarse y medirse su aplicabilidad es la repetición de hechos con esas mismas características, pero no de casos individuales. Por consiguiente, el argumento no es concluyente. Sin embargo, nos brinda la ocasión de destacar un punto importante relacionado con nuestro análisis anterior: cuando hablábamos de la explicación de un hecho singular, el término "hecho" se refería al suceso de cierta característica más o menos compleja en una localización espacio-temporal específica o en un objeto individual determinado, y no a *todas* las características de ese objeto, o a todo lo que ocurre en esa región espacio-tiempo.

Un segundo argumento que debemos mencionar aquí[11] afirma que es imposible establecer generalizaciones científicas —y por ende principios explicativos— acerca de la conducta humana porque las reacciones de un individuo en una situación dada, depende no sólo de esa situación, sino también de los antecedentes personales del individuo. Pero sin duda no hay razón a priori por la cual no puedan lograrse generalizaciones que tomen en cuenta esta dependencia conductal, del pasado del sujeto. Es evidente que en realidad el argumento dado "prueba" demasiado y constituye, por lo tanto, un *non sequitur*, debido a la existencia de ciertos fenómenos físicos, tales como la histéresis magnética y la fatiga elástica, en los cuales la magnitud de un efecto depende de los antecedentes del sistema implicado, y para el cual se han establecido, no obstante, ciertas regularidades generales.

Un tercer argumento insiste en que la explicación de todo fenómeno que implique una conducta intencional exige hacer referencia a motivaciones y, en consecuencia, a un análisis teleológico antes que causal. Por ejemplo, para una enunciación más completa de la explicación sugerida en el caso de la baja de los precios del algodón, habría que indicar las motivaciones del especulador en gran escala como uno de los factores determinantes del hecho en cuestión. Por lo tanto, tenemos que remitirnos a las metas buscadas, lo cual —dice el argumento— introduce un tipo de explicación ajeno a las ciencias físicas. Es cuestionable que muchas de las explicaciones, generalmente incompletas, que se ofrecen para las acciones humanas, implican referencia a propósitos y motivos; pero, ¿acaso esto las hace fundamentalmente distintas de las explicaciones causales de la física y la química? Una diferencia que se sugiere por sí sola reside en la circunstancia de que en la conducta motivada, el futuro parece afectar el presente de una manera que no se encuentra en las explicaciones causales de las ciencias físicas. Pero es evidente que cuando la acción de una persona está motivada, digamos, por el deseo de alcanzar cierto objetivo, no

11 Véase, por ejemplo, la presentación de F. H. Knight sobre este argumento (1924, págs. 251-252).

es el hecho futuro, aún inadvertido, de obtener esa meta lo que determina su conducta presente, puesto que en realidad la meta bien pudiera no alcanzarse nunca; antes bien, digámoslo crudamente, es *a*) su deseo, presente antes de la acción, de alcanzar ese objetivo particular, y *b*) su creencia, también presente antes de la acción, de que tal y cual curso de acción tenga probablemente el efecto deseado. Por consiguiente, los motivos y las creencias determinantes deben clasificarse entre las condiciones antecedentes de una explicación motivacional, y aquí no existe diferencia formal alguna entre la explicación causal y motivacional.

Tampoco constituye una diferencia esencial entre ambas clases de explicación, el hecho de que los motivos sean inaccesibles a la observación directa de un observador exterior, porque los factores determinantes que se aducen en las explicaciones físicas, con mucha frecuencia resultan inaccesibles a la observación directa. Este es el caso, por ejemplo, cuando se señalan cargas eléctricas opuestas para explicar la atracción mutua de dos bolas de metal. La presencia de esas cargas, aunque escapa a la observación directa, se la puede investigar mediante diversas pruebas indirectas, y eso es suficiente para garantizar el carácter empírico del enunciado explicativo. De manera similar, es posible indagar la presencia de ciertas motivaciones solamente por métodos indirectos, lo cual puede incluir referencias a la expresión lingüística del sujeto estudiado, a los deslices de la lengua o de la pluma, etc.; pero hasta tanto estos métodos sean "determinados funcionalmente" con razonable claridad y precisión, no habrá diferencia esencial, en este aspecto, entre la explicación motivacional y la explicación causal, en física.

Un riesgo potencial en la explicación apoyada por motivos reside en el hecho de que el método conduce por sí solo a la fácil construcción de explicaciones *ex post facto* que carecen de fuerza predictiva. Una acción a menudo se explica atribuyéndola a motivos que se conjeturan sólo después que la acción se ha llevado a cabo. Mientras este procedimiento no es de por sí objetable, su solidez requiere que 1) los supuestos motivacionales en cuestión puedan comprobarse, y 2) que se disponga de leyes generales adecuadas que conduzcan al poder explicativo hacia los motivos supuestos. El descuido de estos requisitos con frecuencia priva a la explicación motivacional declarada, de su significación cognitiva.

Algunas veces, la explicación de un acto en función de los motivos del agente se considera como una clase especial de explicación teleológica. Como ya se señaló antes, la explicación motivacional, si se formula adecuadamente, conforma las condiciones de la explicación causal, de modo que el término "teleológica" es inadecuado si se quiere significar ya un carácter no causal de la explicación, ya una determinación peculiar del presente por el futuro. No obstante, si se tiene presente esta condición, el término "teleológico" puede considerarse en ese contexto como referido a explicaciones causales en las cuales algunas de las condiciones antecedentes son motivos del agente cuyos actos habrá que explicar.[12]

[12] Para un análisis lógico detallado del concepto de motivación en la teoría psicológica, véase Koch (1941). El artículo de Rosenblueth, Wiener y Bigelow es una estimulante exposición de la conducta teleológica desde el punto de vista de la física

Las explicaciones teleológicas de esta clase deben diferenciarse de otro tipo de más vasto alcance, que ha sido considerado por ciertas escuelas filosóficas, indispensable especialmente en biología. Consiste en explicar las características de un organismo remitiéndolas a ciertos fines o propósitos, a cuyas características dicen servir. En contraposición a los casos que examinamos antes, no se presume aquí que el organismo persiga los fines, sea consciente o subconscientemente. Así, para explicar el fenómeno del mimetismo, se dice que éste sirve al propósito de proteger al animal con él dotado de ser descubierto por sus perseguidores, y que así tiende a conservar la especie. Antes de que pueda apreciarse la fuerza potencial explicativa de las hipótesis teleológicas de este tipo, debe aclararse su significado. Si de algún modo intentan expresar la idea de que los propósitos a que se refieren son inherentes al plan del universo, entonces es obvio que no pueden comprobarse empíricamente y por lo tanto violan el requisito (R3) de la sección 2. Sin embargo, en ciertos casos, las afirmaciones acerca de los propósitos de las características biológicas pueden trasladarse a enunciados de terminología teleológica que afirmen que esos rasgos funcionan de manera específica, imprescindible, para conservar vivo al organismo o para preservar la especie.[13] El intento de afirmar con exactitud el significado de la aseveración anterior —o la similar de que si no fuera por esas características, dejando invariables las demás circunstancias, el organismo o la especie no sobrevivirían— se enfrenta con dificultades considerables. Pero no necesitamos discutirlas aquí. Porque aunque supusiéramos que los enunciados biológicos de forma teleológica pudieran traducirse adecuadamente en enunciados descriptivos sobre la función conservadora de vida de ciertas características biológicas, es obvio que 1) en estos contextos no es esencial el empleo del concepto de intención, puesto que el término puede eliminarse de esos enunciados por completo, y 2) los supuestos teleológicos, aunque dotados ahora de contenido empírico, no pueden servir como principios explicativos en los contextos comunes. Por ejemplo, el hecho de que determinada especie de mariposa posea un tipo particular de colorido no puede inferirse y, por ende, explicarse a partir de esa enunciación de que ese tipo de color tiene el efecto de proteger las mariposas de las aves que las persiguen ni tampoco puede inferirse la presencia de glóbulos rojos en la sangre humana del enunciado de que tengan la función específica de asimilar oxígeno, que es esencial para la conservación de la vida.

Una de las razones de la perseveración de las consideraciones teleológicas en biología, reside probablemente en lo fructífero del enfoque teleológico como recurso heurístico; indagaciones biológicas que estaban motivadas psicológicamente por una orientación teleológica, por un interés en los objetivos naturales han conducido con frecuencia a importantes resul-

---

y de la biología contemporáneas. La explicación lógica por razones motivacionales se examina con más detalle en el capítulo IV.

[13] En la obra de Woodger (1929), especialmente las págs. 423 y sigs., puede hallarse un análisis de los enunciados teleológicos en biología según estos lineamientos; Kaufmann, en su obra (1944, cap. 8) defiende en esencia, la misma interpretación.

tados que pueden formularse con terminología no teleológica, y que acrecientan nuestro conocimiento científico de las conexiones causales entre los fenómenos biológicos.

Otro aspecto que atrae la atención hacia las consideraciones teleológicas es su carácter antropomórfico. Una explicación teleológica tiende a hacernos sentir que verdaderamente "comprendemos" el fenómeno en cuestión porque está explicado en función de propósitos, con los cuales estamos familiarizados por nuestra propia experiencia de conducta intencional. Pero es importante distinguir aquí entre la comprensión en el sentido psicológico de una sensación de familiaridad empática, y la comprensión en el sentido teórico o cognitivo de exhibir el fenómeno que se debe explicar como un caso especial de cierta regularidad general. La frecuente insistencia en que la explicación significa reducir algo desconocido a ideas o experiencias familiares, conduce por cierto a error. Pues si bien algunas explicaciones científicas tienen este efecto psicológico, en modo alguno es universal: la libre caída de un cuerpo físico puede decirse que es un fenómeno más familiar que la ley de gravedad, mediante la cual puede explicarse; y con toda seguridad las ideas básicas de la teoría de la relatividad resultarán, para muchos, menos familiares que los fenómenos que explican la teoría.

La "familiaridad" del explanans no sólo no es necesaria para una explicación seria como acabamos de señalar, sino que tampoco es suficiente. Esto se demuestra en una cantidad de casos en que el explanans propuesto suena sugestivamente familiar, pero un examen más detenido prueba ser una mera metáfora, o que carece de capacidad para verificarse, o que no incluye leyes generales y, por lo tanto, no tiene poder explicativo. Un ejemplo que viene al caso es el intento neovitalista de explicar los fenómenos biológicos con referencia a una entelequia o fuerza vital. El punto crucial no es aquí cómo se ha pretendido algunas veces el hecho de que las entelequias no puedan verse u observarse directamente; pues también esto es verdad respecto de los campos gravitacionales y, no obstante, es esencial referirse a esos campos para explicar varios fenómenos físicos. La diferencia decisiva entre ambos casos reside en que la explicación física proporciona: 1) métodos de prueba, aunque indirectos, de aseveraciones sobre campos gravitacionales, y 2) leyes generales relacionadas con la fuerza de los campos gravitacionales y la conducta de los objetos que se mueven en ellos. Las explicaciones por medio de entelequias no satisfacen ninguna analogía de estas condiciones. No cumplir con esta primera condición comporta una violación de (R3); hace inaccesibles a la comprobación empírica todos los enunciados sobre entelequias y, en consecuencia, desprovistos de significación empírica. Faltar a la segunda condición implica la violación de (R2). Despoja al concepto de entelequia de todo valor explicativo, porque el poder explicativo nunca reside en un concepto sino en las leyes generales dentro de las cuales funciona. Por consiguiente, no obstante la sensación de familiaridad que evoca, el enfoque neovitalista no puede proporcionar comprensión teórica.

Las observaciones precedentes sobre la familiaridad y la comprensión pueden aplicarse, de manera similar, al criterio sostenido por algunos

pensadores de que la explicación —o la comprensión— de las acciones humanas requiere una comprensión empática de las personalidades de los sujetos.[14] Esta comprensión de otra persona según el propio funcionamiento psicológico puede ser un recurso heurístico útil en la búsqueda de principios psicológicos generales capaces de proporcionar una explicación teórica; pero la existencia de empatía por parte del científico no constituye una condición necesaria ni suficiente para la explicación, o la comprensión científica, de ningún acto humano. No es necesaria, porque algunas veces puede explicarse y predecirse en función de principios generales la conducta de psicóticos o de gente que pertenece a culturas muy diferentes de la del científico, aun cuando aquel que establece o aplica esos principios no pueda entender empáticamente a los sujetos. Y la empatía no es suficiente para garantizar una explicación seria, puesto que puede existir un fuerte sentimiento de empatía aun en ocasiones en que estamos completamente errados al juzgar una personalidad dada. Además, como lo ha señalado Zilsel, la empatía conduce fácilmente a resultados incompatibles; por ejemplo, cuando la población de una ciudad ha estado sometida a bombardeos aéreos pesados durante mucho tiempo, en el sentido empático podemos entender que se haya desmoralizado totalmente; pero con igual facilidad podemos comprender también que haya desarrollado un espíritu de resistencia desafiante. Los argumentos de este tipo a menudo parecen muy convincentes, pero son de carácter *ex post facto* y carecen de significación cognitiva, a menos que se los complete con principios explicativos verificables en forma de leyes o teorías.

En consecuencia, la familiaridad con el explanans, no importa si se logra mediante el uso de una terminología teleológica o de metáforas neovitalistas o por otros medios, no indica el contenido cognitivo ni la fuerza predictiva de una explicación propuesta. Además, el grado en que una idea es considerada familiar varía de una persona a otra y de un momento a otro, y un factor psicológico de este tipo no puede servir de normas para evaluar el mérito de una explicación propuesta. El requisito decisivo para toda explicación sólida es que subsuma el explanandum en leyes generales.

II. SOBRE LA IDEA DE EMERGENCIA

4. *Niveles de la explicación. Análisis de la emergencia*

Según se señaló anteriormente, un fenómeno puede explicarse por un conjunto de leyes de diferentes grados de generalización. Las posiciones cambiantes de un planeta, por ejemplo, pueden explicarse mediante la subsunción en las leyes de Kepler, o por derivación de la ley general de gravedad, que es más inclusiva, en combinación con las leyes de movimiento; o, finalmente, por deducción de la teoría general de la relatividad que ex-

[14] Para un análisis más detallado de este concepto sobre la base de los principios generales que se bosquejaron en el texto, véase Zilsel (1941, secs. 7 y 8) y Hempel (1942, sec. 6).

plica y modifica apenas el conjunto de leyes precedentes. De igual manera, puede explicarse la dilatación de los gases a temperatura creciente y presión constante por medio de la ley de los gases o de la teoría cinética del calor, que es mucho más inclusiva. Esta última explica la ley de los gases y, por lo tanto, de modo indirecto el fenómeno mencionado, por medio de: 1) ciertos supuestos que conciernen a la microconducta de los gases (más específicamente, a las distribuciones de las disposiciones y velocidades de las moléculas gaseosas), y 2) ciertos macro-micro-principios que vinculan las macro-características de un gas, tales como su temperatura, presión y volumen, con las micro-características mencionadas.

Basada en estos ejemplos, con frecuencia se hace una distinción entre los diversos *niveles de explicación.*[15] El primer nivel está representado por la subsunción de un fenómeno bajo una ley general que conecta directamente las características obsevables; los niveles superiores requieren el empleo de constructos teóricos más o menos abstractos que funcionan dentro del contexto de alguna teoría más inclusiva. Como lo demuestran los ejemplos anteriores, el concepto de explicación de nivel superior cubre procedimientos de carácter más bien diferente; uno de los más importantes consiste en explicar una clase de fenómenos mediante una teoría concerniente a su microestructura. Son ejemplos de este método la teoría cinética del calor, la teoría atómica de la materia, la teoría electromagnética y cuántica de la luz y la teoría genética de la herencia. A menudo se cree que sólo el descubrimiento de una microteoría posibilita una real comprensión científica de cualquier tipo de fenómeno porque, para decirlo así, sólo ella nos permite un insight del mecanismo interior del fenómeno. En consecuencia, se ha considerado que las clases de hechos para las que no se disponía de una microteoría, no estaban realmente entendidas; y todo lo relacionado con el status teórico de los fenómenos que carecen de explicación, se señala como raíz de la doctrina de emergencia.

Hablando en términos generales, se ha utilizado el concepto de *emergencia* para caracterizar ciertos fenómenos como "nuevos" y no en el simple sentido psicológico por ser inesperados,[16] sino en el sentido teórico por ser inexplicables e impredecibles de acuerdo con la información referente a las partes espaciales u otros componentes de los sistemas en los cuales ocurren los fenómenos, a los que en este contexto, se denominan con frecuencia "totalidades". Así, por ejemplo, las características del agua tales como su transparencia y estado líquido a la temperatura y presión atmosférica ambientes, o su capacidad para saciar la sed, se han considerado emergentes fundadas en que no hubiese sido posible pronosticarlas a partir del conocimiento de las propiedades de sus componentes químicos, hidrógeno y oxígeno. Por el contrario, se ha dicho que el peso del compuesto no es emergente sino una simple "resultante" de sus componentes y podría haberse pronosticado por una suma sencilla, aun antes de que el compuesto estuviese formado. Las concepciones de explicación y pro-

[15] Para una breve y clara exposición de esta idea, véase Feigl (1945, págs. 284-288).

[16] Con respecto al concepto de novedad en sus significados lógico y psicológico véase también Stace (1939).

nóstico que subyacen a esta idea de emergencia exigen varias observaciones críticas y los cambios correspondientes en el concepto de emergencia.

1. Primero, el problema de que una determinada característica de una "totalidad" *w*, sea o no emergente, no podrá plantearse con sentido hasta que se haya establecido lo que ha de entenderse por parte o componentes de *w*. El volumen de una pared puede inferirse, por ejemplo, de la suma de sus ladrillos, pero no es inferible de los volúmenes de los componentes moleculares de la pared. Por lo tanto, antes de preguntarnos si la característica *W* de un objeto *w* es emergente, deberemos establecer el sentido del término "parte de", mediante la definición de una relación específica *Pt* y la afirmación de que aquellos objetos, y sólo aquellos, que estén en relación *Pt* con respecto a *w* serán partes o componentes de *w*. Puede definirse "*Pt*" como "ladrillo integrante de" (con respecto a compuestos químicos o a un objeto material cualquiera), o "célula de" (con respecto a organismos), etc. Se usará aquí el término "totalidad" con prescindencia de sus diversas connotaciones, y como simple referencia a un objeto cualquiera *w* respecto del cual otros objetos se encuentran en relación *Pt* especificada. Con el objeto de destacar la dependencia del concepto de parte en la definición de la relación *Pt* en cada caso, hablaremos algunas veces de partes *Pt* para referirnos a aquéllas como determinadas por la relación particular *Pt* en consideración.

2. Consideremos ahora un segundo punto crítico. Si se califica de emergente a una característica de una totalidad sencillamente si su aparición no puede inferirse del conocimiento de todas las propiedades de sus partes, entonces, como lo señaló Grelling, ninguna totalidad puede tener característica emergente alguna. Así, y para dar un ejemplo con referencia al anterior, las propiedades del hidrógeno combinado adecuadamente con el oxígeno, incluyen la de formar un compuesto líquido, transparente, etc. De ahí que el estado líquido, la transparencia, etc., del agua puedan inferirse de ciertas propiedades de sus componentes químicos. Por consiguiente, si el concepto de emergencia no resulta vacuo, será necesario especificar en cada caso una clase *G* de atributos y denominar a cierta característica *W* de un objeto *w* como emergente con respecto a *G* y con *Pt* si la ocurrencia de *W* en *w* no puede inferirse de una caracterización completa de todas las partes *Pt* en relación con los atributos contenidos en *G*, es decir, de un enunciado que indique a qué partes de *w* se aplicará cada atributo de *G*. Es evidente que la ocurrencia de una característica puede ser emergente respecto de una clase de atributos y no emergente respecto de otra. Los tipos de atributos que tienen en mente los emergentistas y que por lo general no están indicados de modo explícito, deberán construirse como no triviales, es decir, como que no entrañan lógicamente la propiedad de cada componentes de formar, junto con todos los otros, un todo con las características que se investigan. Algunos casos bastante simples de emergencia, en el sentido hasta aquí especificado, surgen cuando la clase *G* se limita a ciertas propiedades simples de las partes, a la exclusión de las relaciones especiales o de otro tipo que se den entre ellas. Así, no se puede inferir la fuerza electromotriz de un sistema de varias baterías

eléctricas sólo de las fuerzas electromotrices de sus componentes, sin la descripción, en función de conceptos relacionales, de la manera en que las baterías se intercomunican.[17]

3. Por último, la predictibilidad de una determinada característica de un objeto basada en la información específica que concierne a sus partes dependerá, es evidente, de las leyes o teorías generales de que se disponga.[18] Así, el flujo de una corriente eléctrica en un cable que conecta un trozo de cobre con otro de zinc, sumergidos parcialmente en ácido sulfúrico, resulta inexplicable sobre la base de una información concerniente a todo conjunto de atributos no triviales del cobre, del zinc y del ácido sulfúrico, y la estructura particular del sistema que se estudia, a menos que la teoría disponible contenga ciertas leyes generales que se refieran al funcionamiento de las baterías, o aun principios inclusivos de la físico-química. Si, por otro lado, la teoría incluye tales leyes, entonces se puede predecir la presencia de la corriente. Otro ejemplo que al mismo tiempo ilustra el punto 2. tratado más arriba, está dado por la actividad óptica de ciertas sustancias. Por ejemplo, la actividad óptica del ácido sarcoláctico, es decir, el hecho de que su solución haga girar el plano del polarización de la luz planopolarizada, no puede pronosticarse sobre la base de las características químicas de sus elementos componentes; antes hay que conocer ciertos hechos concernientes a las relaciones de los átomos que constituyen la molécula de ácido sarcoláctico. El punto esencial es que esta molécula contiene un átomo de carbono asimétrico, o sea que está formada por cuatro átomos o grupos diferentes, y si se suministra esta información relacional puede pronosticarse la actividad óptica de la solución, siempre y cuando la teoría disponible a tal fin contenga, también, la ley de que la presencia de un átomo asimétrico de carbono en una molécula implica la actividad óptica de la solución; si la teoría no incluye esta micro-macro-ley entonces el fenómeno es emergente respecto de esa teoría.

[17] Esta observación relaciona el presente análisis con un problema básico de la teoría de la Gestalt. Así, por ejemplo, la insistencia en que "el todo es más que la suma de sus partes" puede interpretarse como referido a las características de totalidades cuya predicción requiere el conocimiento de ciertas relaciones estructurales entre las partes. Para un examen más amplio de este punto, véase Grelling y Oppenheim (1937-1938 y 1939).

[18] Grelling y, en forma muy explícita, Henle (1942) propusieron un análisis lógico de la emergencia que hace referencia a las teorías disponibles. En efecto, la definición de Henle caracteriza un fenómeno como emergente cuando no se puede predecir por medio de las teorías aceptadas en el momento, sobre la base de los datos disponibles antes de su aparición. En esta interpretación de la emergencia no se hace referencia a las características de las partes o de los componentes. El concepto de predictabilidad de Henle difiere del implícito en nuestro análisis (explícito en la parte III de este capítulo), en que implica su condición deducible de la hipótesis "más simple" que pueda construirse con los datos y teorías disponibles en ese momento. El artículo de Bergmann (1949) presenta varias observaciones interesantes sobre la idea de emergencia y sobre el análisis de Henle. La idea de que el concepto de emergencia, al menos en algunas de sus aplicaciones, significa referirse a la impredictibilidad mediante leyes "simples", también fue propuesta por Grelling en la correspondencia mencionada en la nota 1. Sin embargo, confiar en la idea de simplicidad de las hipótesis implica dificultades considerables; de hecho, aún no disponemos de ninguna definición satisfactoria de ese concepto.

Se adelanta a veces un argumento referente a que fenómenos tales como el flujo de la corriente o la actividad óptica, según nuestros últimos ejemplos, son absolutamente emergentes, por lo menos en el sentido de que se los hubiera podido predecir, antes de haber sido observados por primera vez; con otras palabras, que no se hubiera logrado formular las leyes que son indispensables para pronosticar tales fenómenos basadas en la información disponible antes de su primera aparición.[19] Sin embargo, este criterio no puede sostenerse. Fundándose en el poder de los datos disponibles en un momento dado, la ciencia establece a menudo generalizaciones mediante las cuales puede pronosticar la ocurrencia de hechos cuyos similares nunca se habían encontrado antes. Así, las generalizaciones fundadas en las periodicidades exhibidas por las características de los elementos químicos conocidos a la sazón, permitieron a Mendeleiev, en 1871, pronosticar la existencia de cierto elemento nuevo y establecer correctamente varias de sus propiedades, así como de algunos de sus compuestos; el elemento en cuestión, el germanio, se descubrió en 1886. Un ejemplo más reciente de este mismo punto está dado por el desarrollo de la bomba atómica y la predicción, basada en principios teóricos establecidos antes de ese hecho, de su explosión en ciertas condiciones específicas y su liberación de una energía devastadora.

Como Grelling subrayó, la observación de que la predictibilidad de la ocurrencia de una característica cualquiera depende del conocimiento teórico disponible, se aplica aun a aquellos casos en los cuales, según el lenguaje de algunos emergentistas, la característica del todo es una simple resultante de las características correspondientes de las partes y puede obtenerse de la suma de estas últimas. Así, hasta el peso de una molécula de agua no puede inferirse de los pesos de sus componentes atómicos sin la ayuda de una ley que exprese el primero como una función matemática específica de aquéllos. De ninguna manera es autoevidente que esta función sea una suma; sí es una generalización empírica, aunque ni siquiera correcta en sentido realmente estricto, como lo demostró la física de la relatividad.

Si no se logra advertir que el problema de la predictibilidad de un fenómeno no puede plantearse de modo significativo a menos que las teorías disponibles para la predicción se hayan especificado, surgirá el equívoco de que ciertos fenómenos tienen una cualidad misteriosa imposible de

[19] C. D. Broad, quien en el capítulo 2 de su libro (1925) relata con sentido crítico y gran claridad el análisis de los puntos esenciales del emergentismo, acentúa la importancia de las "leyes" de composición en la predicción de las características de una totalidad sobre la base de las de sus partes (*op. cit.*, págs. 61 y sigs.); pero acepta el criterio caracterizado en el texto y lo ilustra específicamente mediante la afirmación de que "si queremos conocer las propiedades químicas (y muchas de las físicas) de un compuesto químico, como el cloruro de plata, es absolutamente necesario estudiar las muestras de *ese compuesto particular.* El punto más importante es que sería también inútil estudiar los compuestos químicos en general y comparar sus propiedades con las de sus elementos, con la esperanza de descubrir una *ley general* de composición, por medio de la cual las propiedades de *cualquier* compuesto químico podrían predecirse al conocerse las propiedades de sus elementos por separado" (pág. 64). En el texto se señala que ha sido posible un logro precisamente de esta índole basado en el sistema peródico de los elementos.

explicar, y que su status de emergente debe aceptarse con "natural devoción", según la expresión de C. L. Morgan. Las observaciones presentadas en la discusión anterior despojan a la idea de emergencia de estas connotaciones infundadas: la emergencia de una característica no constituye un rasgo ontológico inherente de algunos fenómenos; más bien es un indicio de la amplitud de nuestro conocimiento en un momento dado. Por lo tanto, no posee un carácter absoluto sino relativo, y lo que hoy es emergente respecto de las teorías disponibles, puede muy bien perder en el futuro la condición emergente.

Las consideraciones precedentes sugieren la siguiente *redefinición* de emergencia: la aparición de una característica *W* en un objeto *w* es emergente en relación con una teoría *T*, una relación de parte *Pt*, y una clase *G* de atributos, siempre que esa ocurrencia no pueda deducirse mediante *T* a partir de una caracterización de las partes *Pt* de *w* respecto de todos los atributos de *G*.

Esta formulación explica el significado de emergencia con respecto a *hechos* de un cierto tipo, tales como la ocurrencia de algunas características *W* en un objeto *w*. Con frecuencia se atribuye emergencia a las *características* antes que a los hechos; este empleo del concepto de emergencia puede interpretarse como sigue: una característica *W* es emergente en relación con *T*, *Pt* y *G* si su ocurrencia en un objeto *cualquiera* es emergente en el sentido indicado.

En lo que concierne a su contenido cognitivo puede interpretarse de manera aproximada la afirmación emergentista de que los fenómenos de la vida son emergentes, como una formulación elíptica de este enunciado: ciertos fenómenos biológicos especificables no pueden explicarse mediante las teorías físico-químicas de hoy, basados en datos referentes a las características físicas y químicas de los componentes atómicos y moleculares de los organismos. De manera semejante, la tesis de un status emergente de la mente debería tomarse para afirmar que las teorías físicas, químicas y biológicas de hoy no son suficientes para explicar todos los fenómenos psicológicos fundados en los datos referidos a las características físicas, químicas y biológicas de las células o de las moléculas o átomos que forman los organismos estudiados. Pero en esta interpretación, la naturaleza emergente de los fenómenos biológicos y psicológicos se convierte en trivial, pues la descripción de varios fenómenos biológicos requiere términos que no pertenecen al vocabulario de la física y de la química contemporáneas; entonces no podemos esperar que todos los fenómenos específicamente biológicos sean explicables, es decir, deductivamente inferibles, por medio de las teorías físico-químicas disponibles, según las condiciones iniciales que a su vez se describen con la terminología exclusivamente físico-química. Con el objeto de obtener una interpretación menos trivial de la afirmación de que los fenómenos de la vida son emergentes, debemos incluir, por ende, en la teoría explicativa todas aquellas leyes presuntivas que pronto se han de aceptar y que vinculan el "nivel" físico-químico con el biológico; en otras palabras, que contiene, por un lado ciertos términos físicos y químicos, incluso aquellos requeridos para describir estructuras moleculares; y por otro lado, ciertos conceptos de biología. Una observación análoga se

aplica al caso de la psicología. Si se interpreta en este sentido la aseveración de que la vida y la mente tienen un status emergente, entonces su significación puede sintetizarse de modo aproximado en el enunciado de que en función de teorías microestructurales no se dispone hasta el momento de explicación alguna para vastas clases de fenómenos estudiados en biología y psicología.[20]

Las afirmaciones de este tipo, entonces, parecen representar el núcleo racional de la doctrina de la emergencia. En su forma revisada, la idea de emergencia ya no lleva consigo la connotación de absoluta impredictabilidad —noción que es objetable no sólo porque implica y perpetúa ciertos errores lógicos, sino también porque promueve y de manera similar a las ideas del neovitalismo, una actitud de resignación que asfixia la investigación científica. Sin duda esta característica, agregada a la esterilidad teórica, explica el rechazo de la doctrina clásica y absolutista de la emergencia por parte de la mayoría de los científicos contemporáneos.[21]

III. ANALISIS LOGICO DE LA LEY Y LA EXPLICACION

5. *Problemas del concepto de ley general*

De nuestra revisión general de las características de la explicación científica, emprendemos ahora un examen más detenido de la estructura lógica. La explicación de un fenómeno, señalamos, consiste en su subsunción dentro de leyes o una teoría. Pero ¿qué es una ley? ¿Qué es una teoría? Mientras el significado de estos conceptos parece intuitivamente claro, el intento de construir definiciones explícitas y adecuadas se obstaculiza de manera considerable. En esta sección describiremos y analizaremos algunos problemas básicos del concepto de ley; en la siguiente, intentaremos proponer, sobre la base de las sugerencias así obtenidas, definiciones de la ley y de la explicación para un lenguaje formalizado modelo, de una estructura lógica simple.

Aquí interpretaremos el concepto de ley de manera que se aplique sólo a enunciados verdaderos. El procedimiento alternativo aparentemente plausible de requerir un alto grado de confirmación antes que la verdad de

[20] El siguiente pasaje tomado de la obra de Tolman (1932) puede servir para apoyar esta interpretación: "...'los actos conductales', aunque sin duda en absoluta correspondencia unívoca con los hechos moleculares subyacentes de la física y de la fisiología, tienen como totalidades 'molares' ciertas propiedades emergentes propias... Además, estas propiedades molares de los actos de conducta, en el estado actual de nuestros conocimientos —es decir, previamente a la construcción de muchas correlaciones empíricas entre la conducta y sus correlatos fisiológicos— no pueden ni siquiera conocerse por inferencia desde un mero conocimiento de los hechos —moleculares— subyacentes de la física y la fisiología" (*op. cit.*, págs. 7 y 8). De manera similar Hull utiliza la distinción entre teorías molares y moleculares, y señala que actualmente no dispone la psicología de teorías de este último tipo. Véase 1943a, págs. 19 y sigs.; 1943, pág. 275.

[21] Esta actitud del hombre de ciencia es expresada, por ejemplo, por Hull, 1943a, págs. 24-28.

una ley nos parece inadecuado: conduciría a un concepto relativizado de ley que podría expresarse con la frase "la oración S es una ley relacionada con los elementos de prueba *E*". Esto concuerda con la significación que se asigna habitualmente, tanto en la ciencia como en la investigación metodológica, al concepto de ley. Así, por ejemplo, no diríamos que la fórmula general de Bode para la distancia de los planetas desde el sol era una ley relacionada con las pruebas astronómicas disponibles hacia 1770, cuando Bode la propuso, y que perdió su carácter de ley después del descubrimiento de Neptuno y la determinación de su distancia al sol; antes podríamos decir que las pruebas limitadas originales habían señalado una gran probabilidad de certeza a la suposición de que la fórmula era una ley, mientras que una información adicional posterior redujo tanto aquella probabilidad como para que se tuviese prácticamente la seguridad de que la fórmula de Bode no era verdadera en sentido general, y por lo tanto no era una ley.[22]

Además de ser verdadera, una ley debe cumplir una cierta cantidad de condiciones adicionales que pueden estudiarse de modo independiente del requisito fáctico de verdad, porque se refieren, por así decirlo, a todas las leyes lógicamente posibles, sean fácticamente verdaderas o falsas. Adoptando un término propuesto por Goodman[23] diremos que una oración *parece legal* si posee todas las características de una ley general con la posible excepción de la verdad. De ahí que toda ley sea una oración legal, pero no viceversa.

Nuestro problema de analizar la noción de la ley se reduce así al de explicar el concepto de esta "oración legal". Estableceremos la clase de estas oraciones legales diciendo que incluyen enunciados analíticos generales, tales como "una rosa es una rosa", así como las oraciones legales de la ciencia empírica, que tienen contenido empírico.[24] No será necesario el requisito de que toda oración legal permisible en contextos explicativos deba pertenecer al segundo tipo; en su lugar, nuestra definición de explicación será construida de modo que garantice el carácter fáctico de la totalidad de las leyes —aunque no de cada una de ellas por separado— que funcionan en la explicación de un hecho empírico.

¿Cuáles son las características de las oraciones legales? En primer término, estas oraciones son enunciados de forma universal, tales como "todos los huevos de petirrojo son de color azul verdoso", "todos los metales son conductores de la electricidad", "a presión constante, todo gas se

[22] La exigencia de que las leyes sean verdaderas trae como consecuencia que jamás pudo conocerse definitivamente si un enunciado empírico dado *E* es una ley; puesto que la oración que afirma la verdad de *E* es lógicamente equivalente a *E* y, por lo tanto, capaz sólo de adquirir una probabilidad, o grado de confirmación más o menos elevado, respecto de las pruebas experimentales disponibles en un momento determinado. Sobre el tema véase Carnap (1946). Para una excelente exposición no técnica del concepto semántico de verdad, que aquí se invoca, se remite al lector a Tarski (1944).

[23] 1947, pág. 125.

[24] Este procedimiento fue sugerido por el enfoque de Goodman en 1947. De manera similar, Reichenbach en un examen detallado del concepto de ley, construye su concepto de enunciado nomológico de modo que incluye tanto las oraciones analíticas como sintéticas; véase 1947, cap. 8.

dilata con el aumento de temperatura". Como estos ejemplos lo ilustran, la oración legal no sólo tiene forma universal sino también condicional; formula una afirmación universal en el sentido de que si cumple un conjunto de condiciones, *C*, se cumplirá también otro conjunto específicamente de condiciones, *E*. La forma estándar de la expresión simbólica de una oración legal es, por ende, un condicional universal. Sin embargo, puesto que todo enunciado condicional puede transformarse en otro no condicional no se considerará esencial para la oración legal, mientras que será indispensable mantener el carácter universal.

Pero el requisito de la forma universal no es suficiente para caracterizar las oraciones legales. Supongamos, por ejemplo, que una cesta *b* contenga en cierto momento *t* una cantidad de manzanas rojas y nada más.[25] Entonces, el enunciado

(E_1) Toda manzana de la cesta *b* en el momento *t* es roja

es verdadero y a la vez de forma universal. No obstante, esta oración no está calificada como una ley; rehusaríamos, por ejemplo, explicar por medio de la subsunción del hecho de que una manzana determinada, elegida al azar en la cesta, sea roja. ¿Qué diferencia a E_1 de una oración legal? Dos puntos, que trataremos de manera sucesiva, se sugieren por sí mismos: el alcance finito y la referencia a un objeto especificado.

Primero, la oración E_1 formula, en efecto, una afirmación respecto de un número finito de objetos, y esto parece irreconciliable con la pretensión de universalidad que se asocia comúnmente a la noción de ley.[26] Pero, ¿no se consideran "legales" las leyes de Kepler aunque se refieran sólo a un conjunto finito de planetas? Y ¿no estaríamos dispuestos a considerar como oración legal a la siguiente?:

(E_2) La totalidad de los 16 cubos de hielo de la bandeja del refrigerador tienen una temperatura inferior a 10 grados centígrados.

Podemos conceder este punto; pero existe una diferencia esencial entre E_1, por un lado, y las leyes de Kepler y también E_2, por el otro; sabemos que estas últimas, si bien de extensión finita, son consecuencia de leyes más inclusivas cuyo alcance no es limitado, mientras que éste no es el caso de E_1.

Adoptando un procedimiento sugerido recientemente por Reichenbach[27] distinguiremos, por lo tanto, las leyes fundamentales de las deriva-

[25] La dificultad que ilustra este ejemplo fue enunciada concisamente por Langford (1941), quien se refirió a ella como el problema de distinguir entre universales de hecho y universales causales. Para un análisis e ilustración más amplios de este punto, véase también Chisholm (1946), especialmente las págs. 301 y sigs. Goodman realizó un análisis sistemático del problema (1947, parte III). Si bien no concierne al tema específico en discusión, el examen detallado de los condicionales contrafácticos y sus relaciones con las leyes de la naturaleza, en el capítulo 8 de la obra de Lewis (1946), contiene importantes observaciones sobre varios de los problemas planteados en la presente sección.

[26] Popper (1935, sec. 13) y Reichenbach (1947, pág. 369) expresaron el criterio de que las leyes deben interpretarse como no limitadas a un ámbito finito.

[27] 1947, pág. 361. No obstante, nuestra terminología así como las definiciones

das. Denominaremos ley derivada a un enunciado que tenga carácter universal y se origine de ciertas leyes fundamentales. El concepto de ley fundamental requiere mayor esclarecimiento; por ahora, diremos que las leyes fundamentales deben satisfacer una determinada condición: alcance ilimitado.

Sería excesivo, no obstante, negar el status fundamental de la oración legal a todos los enunciados que afirman, en efecto, sólo acerca de una clase finita de objetos, porque ello excluiría una oración tal como "todos los huevos de petirrojo son de color azul verdoso", puesto que, presumiblemente, la clase de todos los huevos de petirrojo (pasados, presentes y futuros) es finita. Pero otra vez vemos aquí una diferencia esencial entre esta oración y, digamos, E_1. Para establecer la finitud de la clase de los huevos de petirrojo se requiere conocimiento empírico, mientras que si se concibe la oración E_1 de manera intuitivamente ilegal, los términos "cesta *b*" y "manzana" se comprenden como que implican la finitud de la clase de las manzanas que están en la cesta en el momento *t*. De esta manera, y por así decirlo, el significado de sus términos constitutivos, por sí solo —sin información fáctica adicional—, implica que E_1 tiene extensión finita. Las leyes fundamentales, entonces, deberán interpretarse de modo que satisfagan la condición de alcance no limitado; nuestra formulación de esta condición, sin embargo, que remite a lo que está implicado en "el significado" de ciertas expresiones, es demasiado vaga y deberá ser reexaminada más adelante. De paso señalamos que la estipulación referida también excluye de la clase de las oraciones legales fundamentales a candidatos tan indeseables como "Todos los objetos uránicos son esféricos", en que "uránico" designa la propiedad de ser el planeta Urano; por cierto que, aunque posee forma universal, esta oración no satisface la condición de alcance no limitado.

En nuestra búsqueda de una caracterización general de las oraciones legales, tomaremos ahora en cuenta un segundo indicio que proporciona la oración E_1. Además de violar la condición de extensión no limitada, tiene la peculiaridad de referirse a un objeto particular, la cesta *b*; y al parecer, esto ambién viola el carácter universal de una ley.[28] La restricción que parece indicarse aquí, debiera aplicarse nuevamente sólo a las oraciones legales fundamentales, porque un enunciado general verdadero sobre la caída libre de los cuerpos físicos en la Luna, si bien se refiere a un objeto particular, constituiría aun una ley, aunque fuese derivada.

En consecuencia, parece razonable estipular que una oración legal fundamental debe ser de forma universal y no ha de contener ocurrencias esenciales (es decir, no eliminables) de designaciones de objetos particulares. Pero esto no es suficiente; realmente en este punto se presenta una dificultad especialmente seria. Considérese la oración

---

que se propondrán más adelante para los dos tipos de ley no coinciden con las de Reichenbach.

[28] En física, la idea de que una ley no debe referirse a ningún objeto particular ha hallado su expresión en la máxima de que las leyes generales de la física no deben contener referencia a puntos espacio-temporales específicos, y que las coordenadas espacio-temporales deben ocurrir en ellas solamente en la forma de diferencias o diferenciales.

(E_3) Todo lo que sea una manzana de la cesta b en el momento t o una muestra de óxido férrico, será rojo.

Si empleamos una expresión especial, por ejemplo "x es una manfer" como sinónimo de "x es o bien una manzana de la cesta b en t o bien una muestra de óxido férrico", entonces el contenido de E_3 puede expresarse como sigue:

(E_4) Todo lo que es "manfer" es rojo.

El enunciado que así se obtiene es de forma universal y no contiene designaciones de objetos particulares, e igualmente satisface la condición de alcance ilimitado; pero sin embargo es evidente que E_4 no puede calificarse como una oración legal fundamental del mismo modo que E_3.

Mientras "manfer" sea un término definido en nuestro lenguaje, puede satisfacer la dificultad estipulando que después de haber eliminado los términos definidos, una oración legal fundamental no debe contener ocurrencias esenciales de designaciones de objetos particulares. Pero esta salida no sirve cuando "manfer", u otro término del mismo tipo, constituye un predicado primitivo del lenguaje estudiado. Esta reflexión indica que hay que imponer ciertas restricciones a los predicados (términos de propiedades o relaciones) que pueden presentarse en las oraciones legales fundamentales.[29]

De manera más específica, la idea sugiere admitir un predicado en una oración legal fundamental —sólo cuando sea universal puro— o, digamos, de índole cualitativo puro; expresado de distinto modo, si un enunciado de su significado no exige hacer referencia a objeto concreto alguno ni a ninguna localización espacio-temporal. Así, los términos "suave", "verde", "más caliente que", "mientras", "líquido", "con carga eléctrica", "mujer", "padre de", constituyen predicados cualitativos puros, mientras que "más alto que la Torre Eiffel", "medieval", "lunar", "ártico", "Ming" no lo son.[30]

[29] El punto ilustrado por las oraciones E_3 y E_4 fue señalado por Goodman, quien también destacó la necesidad de imponer ciertas restricciones a los predicados cuya ocurrencia pueda ser admisible en las oraciones legales. Estos predicados son esencialmente los mismos que Goodman llama proyectables. Goodman ha sugerido que los problemas para establecer criterios precisos de proyectabilidad, para interpretar condicionales contrafácticos y para definir el concepto de ley están tan íntimamente relacionados, que son virtualmente aspectos de un único problema. (Véase sus artículos, 1946 y 1947.) Carnap hizo una sugerencia para el análisis de la proyectabilidad en 1947. La nota de Goodman (1947a) contiene observaciones críticas sobre las propuestas de Carnap.

[30] Popper argumentó (1935, secs. 14 y 15) que las leyes, además de ser de forma universal, deben contener sólo predicados universales puros. Nuestra expresión "predicado cualitativo puro", de índole alternativa, se eligió por analogía con el término de Carnap "propiedad cualitativa pura" (véase 1947). La caracterización dada en el texto de los predicados universales puros parece preferible a una más simple y tal vez más común, para que el enunciado del significado del predicado no tenga que referirse a objetos particulares. Esta formulación podría ser demasiado excluyente, puesto que podría argüirse que enunciar el significado de términos cualitativos puros, tales como "azul" o "caliente", requiere tomar como ejemplo algún

Excluir de las oraciones legales fundamentales los predicados que no sean cualitativos puros, aseguraría al mismo tiempo el cumplimiento de la condición de alcance ilimitado, porque el significado de un predicado cualitativo puro no requiere una extensión finita; y, en realidad, todas las oraciones consideradas antes que violan la condición de alcance ilimitado hacen referencia explícita o implícita a objetos específicos.

No obstante, la estipulación que acabamos de proponer sufre de vaguedad del concepto de predicado cualitativo puro. La pregunta de si indicar el significado de un predicado en inglés exige o no hacer referencia a algún objeto específico, no siempre permite una respuesta inequívoca, puesto que el inglés, siendo un lenguaje natural, no proporciona definiciones explícitas u otras explicaciones claras de significado con respecto a sus términos. Por consiguiente, parece razonable intentar definir el concepto de ley, no con respecto al inglés o a cualquier otro idioma natural, sino más bien con respecto a un lenguaje formalizado —llamémoslo lenguaje modelo *L*— gobernado por un sistema bien determinado de reglas lógicas, y en el cual todo término está caracterizado como primitivo o introducido por medio de una definición explícita en función de los primitivos.

Esta referencia a un sistema bien determinado es habitual en la investigación lógica y, por cierto, muy natural en el contexto de todo intento para desarrollar criterios precisos para ciertas distinciones lógicas. Pero por sí solo no es suficiente para vencer la dificultad específica que se analiza. Pues si bien ahora es posible caracterizar fácilmente como no cualitativos puros todos aquellos predicados entre los definidos en *L*, cuyo definiens contenga una ocurrencia esencial de algún nombre individual, nuestro problema aún no está resuelto para los términos primitivos del lenguaje, cuyos significados no están determinados por definiciones dentro del lenguaje sino más bien por reglas semánticas de interpretación. Porque queremos permitir la interpretación de los primitivos de *L* mediante atributos tales como azul, duro, sólido, más caliente que, etc., pero no por medio de la propiedad de ser un descendiente de Napoleón, un animal ártico o una estatua griega. Y la dificultad reside precisamente en establecer criterios rigurosos para diferenciar las interpretaciones admisibles y las no admisibles. Así se plantea nuevamente el problema de encontrar una definición adecuada para los atributos cualitativos puros, es decir, para los conceptos del metalenguaje en que se formula la interpretación semántica de los términos primitivos. Podemos postergar enfrentarnos con esa dificultad mediante la suposición de la calidad formal del metalenguaje semántico, del metalenguaje, etc., pero en algún punto deberemos detenernos ante un metalenguaje no formalizado, y para él se necesitará una caracterización de los predicados cualitativos puros que presentará en alto grado los mismos problemas que los del idioma inglés no formalizado con

---

objeto particular que tenga la cualidad señalada. El punto esencial es que no es necesario elegir ningún objeto específico; cada uno de los objetos azules o calientes del conjunto lógicamente ilimitado puede servir. Sin embargo, al explicar el significado de "más alto que la torre Eiffel", "ser una manzana de la cesta en el momento *t*", "medieval", etc., hay que hacer referencia a un objeto específico o a alguno que pertenezca a un conjunto limitado de objetos.

el cual habíamos comenzado. La caracterización de un predicado cualitativo puro como aquel cuyo significado puede hacerse explícito sin referencia a ningún objeto concreto señala el significado que se intenta pero no lo explica con precisión, y queda en pie el problema de una definición adecuada de los predicados cualitativos puros.

Sin embargo, no cabe duda de que existe una gran cantidad de predicados que generalmente se reconocerían más bien como cualitativos puros en el sentido aquí señalado, y como admisibles en la formulación de oraciones legales fundamentales; ya hemos dado algunos ejemplos y la lista podría ampliarse con facilidad. En adelante, cuando hablemos de predicados cualitativos puros, tendremos en mente los predicados de este tipo.

En la sección siguiente se describirá un lenguaje modelo *L* de estructura lógica más bien simple, cuyos términos primitivos se supondrán cualitativos en el sentido ya indicado. Se definirán los conceptos de ley y explicación para este lenguaje, de modo que se tomen en cuenta las observaciones generales expuestas en la presente sección.

6. *Definición de ley y explicación en un lenguaje modelo*

Respecto de la sintaxis de nuestro lenguaje modelo *L* señalaremos los siguientes supuestos: *L* tiene la estructura sintáctica del cálculo funcional inferior, sin signo de identidad. Además de los signos de negación, alternancia (disyunción), conjunción e implicación (condicional), y de los símbolos de cuantificación universal y existencial respecto de variables individuales, el vocabulario de *L* posee constantes individuales ("*a*", "*b*", . . .), variables individuales ("*x*", "*y*", . . .) y predicados de cualquier grado finito que se desee; estos últimos pueden incluir, en especial, predicados de grado 1 ("*P*", "*Q*", . . .) que expresan propiedades de individuos, y predicados de grado 2 ("*R*", "*S*", . . .) que expresan relaciones diádicas entre individuos.

Para mayor simplicidad suponemos que todos los predicados son primitivos, es decir, indefinidos en *L*, o bien que antes de que los criterios que se desarrollarán a continuación se apliquen a una oración, todos los predicados definidos que ésta contenga serán eliminados en favor de los primitivos.

Las reglas sintácticas de *L* para la formación de oraciones y para la inferencia lógica son las del cálculo funcional inferior. Ninguna oración puede contener variables libres, de modo que la generalidad se exprese siempre mediante cuantificación universal.

Para referencia posterior, definiremos ahora, en términos puramente sintácticos, algunos conceptos auxiliares. En las definiciones siguientes, *E* es entendida siempre como una oración.

(*6.1a*) *E* es formalmente verdadera (formalmente falsa) dentro de *L* si *E* (la negación de *E*) puede probarse dentro de *L*, es decir, mediante las reglas formales de la inferencia lógica para *L*. Si dos oraciones de *L* pueden deducirse recíprocamente se las denominará equivalentes.

(*6.1b*) Se dice que *E* es una oración singular o, alternativamente,

molecular si no contiene variables. Una oración singular que no contiene enunciados conectivos se denomina también atómica. Ejemplos: las oraciones "$R(a,b) \supset [P(a) \cdot \sim Q(a)]$", "$\sim Q(a)$", "$R(a,b)$", "$P(a)$" son todas singulares o moleculares; las dos últimas son atómicas.

(*6.1c*) Se dice que *E* es una oración generalizada si consiste de uno o más cuantificadores seguidos por una expresión que no contiene cuantificadores. Se dice que *E* tiene forma universal si es una oración generalizada y todos los cuantificadores que contiene son universales. *E* es llamada generalizada pura (universal pura) si *E* es una oración (tiene forma universal) y no contiene constantes individuales. Se dice que *E* es esencialmente universal y no equivale a una oración singular. *E* se denomina esencialmente generalizada si es generalizada y no equivale a una oración singular.

Ejemplos: "$(x)[P(x) \supset Q(x)]$", "$(x)R(a,x)$", "$(x)[P(x) \vee P(a)]$", "$(x)[P(x) \vee \sim P(x)]$", "$(Ex)[P(x) \cdot \sim Q(x)]$", "$(Ex)(y)[R(a,x) \cdot \cdot S(a,y)$"]

son todas oraciones generalizadas; las cuatro primeras tienen forma universal, la primera y la cuarta son universales puras; la primera y la segunda son esencialmente universales; la tercera es equivalente a la oración singular "$P(a)$", y la cuarta, a "$P(a) \vee \sim P(a)$". Todas las oraciones, excepto la tercera y la cuarta, son esencialmente generalizadas.

Con respecto a la interpretación semántica de *L*, establecemos las dos estipulaciones siguientes:

(*6.2a*) Todos los predicados primitivos de *L* son cualitativos puros.

(*6.2b*) El universo del discurso de *L*, es decir, el dominio de los objetos cubiertos por los cuantificadores, consta de todos los objetos físicos, o de todas las localizaciones espacio-temporales.

Una estructura lingüística de este tipo no es suficiente para la formulación de teorías científicas, puesto que no contiene "functores" y no proporciona medios para tratar los números reales. Además, subsiste actualmente el problema de si puede construirse un sistema constitutivo en el que todos los conceptos de la ciencia empírica sean reducidos, mediante cadenas de definiciones explícitas, a una base de primitivos de carácter cualitativo puro. Sin embargo, consideramos que vale la pena estudiar estos problemas por el tipo simplificado de lenguaje recién descubierto, porque el análisis de la ley y de la explicación está lejos de ser trivial aun en nuestro lenguaje modelo *L*, y porque dicho análisis también esclarece el carácter lógico de los conceptos que se investigan para su aplicación a contextos más complejos.

De acuerdo con las consideraciones desarrolladas en la sección 5, ahora definiremos:

(*6.3a*) *E* es una oración legal fundamental de *L* si *E* es esencialmente universal pura; *E* es una ley fundamental de *L* si *E* es universal pura y verdadera.

(*6.3b*) *E* es una ley derivada en el lenguaje *L*: 1) si *E* es esencialmente universal, pero no universal pura, y 2) si existe un conjunto de leyes fundamentales de *L* de las cuales *E* es una consecuencia.

(*6.3c*) *E* es una ley de *L* si es una ley fundamental o derivada en el lenguaje *L*.

Es obvio que las leyes fundamentales según se han definido aquí incluyen, además de enunciados generales de carácter empírico, todos aquellos enunciados de forma universal pura que son verdaderos por razones puramente lógicas; es decir, aquellos que son formalmente verdaderos en *L*, como "$(x)[P(x) \vee \sim P(x)]$", y aquellos cuya verdad deriva de manera exclusiva de la interpretación dada a sus componentes, como en el caso de

$$\text{“}(x)[P(x) \supset Q(x)]\text{”}$$

si se interpreta a "*P*" como la propiedad de ser padre, y a "*Q*" la de ser varón. Las leyes derivadas, por otra parte, no incluyen ninguna de estas categorías; por cierto, ninguna ley fundamental es también derivada.[31]

Puesto que los primitivos de *L* son cualitativos puros, todos los enunciados de forma universal de *L* satisfacen también el requisito de alcance no limitado, y así se ve con facilidad que el concepto de ley, tal como lo hemos definido, satisface todas las condiciones sugeridas en la sección 5.[32]

La explicación de un fenómeno puede implicar oraciones generalizadas que no sean de forma universal. Utilizaremos el término "teoría" para referirnos a ellas, y lo definiremos mediante la siguiente cadena de definiciones:

(*6.4a*) *E* es una teoría fundamental si *E* es generalizada pura y verdadera.

(*6.4b*) *E* es una teoría derivada en el lenguaje *L* si: 1) *E* es esencialmente generalizada, pero no generalizada pura, y 2) si existe un conjunto de teorías fundamentales en *L* de las cuales *E* es una consecuencia.

(*6.4c*) *E* es una teoría en *L* si es una teoría fundamental o derivada en *L*.

En virtud de estas definiciones, toda ley es también una teoría, y toda teoría es verdadera.

Con ayuda de los conceptos así definidos, reformularemos ahora con mayor precisión nuestra primera caracterización de la explicación científica con la referencia específica a nuestro lenguaje modelo *L*. Será conveniente establecer nuestros criterios para una explicación seria en la forma

[31] Tal como las hemos definido, las leyes fundamentales incluyen enunciados condicionales universales con antecedentes vacuos, por ejemplo "todas las sirenas son morenas". Este punto no parece conducir a consecuencias inconvenientes en la definición de explicación que se propondrá más adelante. Para un análisis detallado de los condicionales universales con antecedentes vacuos véase Reichenbach (1947, cap. 8).

[32] (Agregado en 1964.) Sin embargo, Nagel ha demostrado que nuestra definición del concepto de ley fundamental es demasiado limitada; véase el postscriptum a este capítulo.

de una definición de la expresión "el par ordenado de oraciones (T, C) constituye un explanans para la oración E". Nuestro análisis estará limitado a la explicación de hechos particulares, es decir, al caso donde el explanandum, E, es una oración singular.[33]

Por analogía con el concepto de oración legal, que no requiere satisfacer el requisito de verdad, introduciremos primeramente el concepto auxiliar de explanans potencial que no está sujeto al requisito de verdad; luego se definirá la noción de explanans con la ayuda de este concepto auxiliar. Las consideraciones presentadas en la parte I sugieren las siguientes estipulaciones iniciales:

(6.5) Un par ordenado de oraciones (T, C) constituye un explanans potencial para una oración singular E sólo si

1) T es esencialmente generalizada y C es singular.

2) E es derivable, en L, de T y C juntas, pero no de C sola.

(6.6) Un par ordenado de oraciones (T, C) constituye un explanans para una oración singular E si y sólo si

1) (T, C) es un explanans potencial para E.

2) T es una teoría y C es verdadera.

(6.6) es una definición explícita de explicación en función del concepto de explicación potencial.[34] Por otra parte, (6.5) no se sugiere como definición, sino como una enunciación de las condiciones necesarias de la explicación potencial. Demostraremos luego que esas condiciones no son suficientes y analizaremos los requisitos adicionales que deben complementar a (6.5) con el objeto de proporcionar una definición de la explicación potencial.

Antes de volver sobre este punto, son necesarias algunas observaciones

[33] No se trata de una libre elección: la reconstrucción racional precisa de la explicación aplicada a las regularidades generales presenta problemas singulares para los cuales no podemos, actualmente, ofrecer ninguna solución. El núcleo de la dificultad se puede indicar brevemente haciendo referencia a un ejemplo: las leyes de Kepler, K, en conjunción con la ley de Boyle, B, forman una ley más fuerte, $K.B.$, pero el derivado de K a partir de esta última ley no sería considerado como explicación de las regularidades enunciadas en las Leyes de Kepler; se pensaría, más bien, que representa de hecho una "explicación" irrelevante de las leyes de Kepler a ellas mismas. Se reconocerá, por el contrario, la derivación de las leyes de Kepler a partir de las leyes de Newton del movimiento y de la gravedad como una genuina explicación en función de regularidades más inclusivas, o sea las llamadas leyes de nivel superior. El problema surge, en consecuencia, de establecer criterios bien definidos para la distinción de niveles de explicación o para comparar oraciones generalizadas en cuanto a su inclusividad. La formulación de criterios adecuados para este propósito es un problema aún no resuelto.

[34] En (6.6)(2) es necesario estipular que T es una teoría y no simplemente que sea verdadera, porque según se ha visto en la sección 5, las oraciones generalizadas que ocurren en un explanans deben constituir una teoría y no toda oración esencialmente generalizada que sea verdadera es realmente una teoría, es decir, una consecuencia de un conjunto de oraciones generalizadas puras verdaderas.

con respecto a la formulación de (6.5). El análisis presentado en la parte I sugiere que un explanans para una oración singular consiste en una clase de oraciones generalizadas y otra de singulares. En (6.5) suponemos que los elementos de cada una de esas clases separadamente, están unidos a una oración. Esta estipulación simplificará nuestras formulaciones y, en el caso de las oraciones generalizadas, tiene un propósito adicional: una clase de oraciones esencialmente generalizadas puede ser equivalente a una oración singular; así, la clase ["$P(a)$ v $(x)Q(x)$", "$P(a)$ v $\sim$ $(x)Q(x)$"] es equivalente a la oración "$P(a)$". Puesto que la explicación científica utiliza de modo esencial las oraciones generalizadas, deben excluirse los conjuntos de leyes de esta naturaleza; esto se logró antes combinando todas las oraciones generalizadas del explanans en una sola conjunción, T, y estipulando que T debe ser esencialmente generalizada. Nuevamente, puesto que la explicación científica utiliza de manera esencial las oraciones generalizadas, E no debe ser consecuencia de C sola: la ley de gravedad combinada con la oración singular "María es rubia y de ojos azules", no constituye un explanans para "María es rubia". La última estipulación en (6.5) introduce la restricción requerida, y prohibe así la autoexplicación completa del explanandum, es decir, la derivación de E de una oración singular que tenga a E por consecuencia. Esa misma restricción elimina también la necesidad de un requisito especial en el sentido de que T debe tener un contenido fáctico si (T, C) ha de ser un explanans potencial de una oración empírica E. Pues si E es fáctica, entonces como E es una consecuencia de T y C unidas y no de C sola, T también debe ser fáctica.

Sin embargo, nuestras estipulaciones en (6.5) no excluyen lo que podría denominarse autoexplicación parcial del explanandum. Considérense las oraciones T_1 = "$(x)[P(x) \supset Q(x)]$", C_1 = "$R(a, b) \cdot P(a)$", E_1 = "$Q(a) \cdot R(a, b)$". Estas oraciones satisfacen todos los requisitos establecidos en (6.5), pero no parece intuitivamente adecuado decir que (T_1, C_1) explique potencialmente a E_1, porque la ocurrencia del componente "$R(a, b)$" de C_1 en la oración E_1 importa una explicación parcial del explanandum por sí misma. ¿No sería posible excluir, mediante una estipulación adicional, todos aquellos casos en que E comparte parcialmente su contenido con C, es decir, donde C y E tienen una consecuencia común que no es formalmente verdadera en L? Esta estipulación sería equivalente al requisito de que C y E tienen que ser alternativamente exhaustivas, en el sentido de que su disyunción es formalmente verdadera pues el contenido que dos oraciones cualesquiera tienen en común se expresa mediante su disyunción. Pero la restricción propuesta, sin embargo, sería demasiado severa, ya que si E no comparte por lo menos algo de su contenido con C, entonces C es completamente innecesaria para la derivación de E a partir de T y C, es decir, E puede inferirse de T sola. Por consiguiente, en toda explicación potencial en que no se pueda prescindir del componente singular del explanans, el explanandum está aplicado parcialmente por sí mismo. Tómese, por ejemplo, la explicación potencial de E_2 = "$Q(a)$" mediante T_2 = "$(x)[P(x) \supset Q(x)]$" y C_2 = "$P(a)$", que satisface a (6.5) y que con seguridad es intuitivamente inobjetable. Sus tres componentes pueden expresarse de modo equivalente por las siguientes oraciones:

$T'_2 =$ "$(x) [\sim P(x) \vee Q(x)]$";
$C'_2 =$ "$[P(a) \vee Q(a)] \cdot [P(a) \vee \sim Q(a)]$";
$E'_2 =$ "$[P(a) \vee Q(a)] \cdot [\sim P(a) \vee Q(a)]$".

Esta reformulación demuestra que parte del contenido del explanandum está incluida en el contenido del componente singular del explanans y, en este sentido, se explica por sí mismo.

Nuestro análisis ha llegado aquí a un punto en que la idea corriente e intuitiva de explicación es demasiado vaga para proporcionar más orientación hacia una reconstrucción racional. En realidad, el último ejemplo sugiere de modo muy claro que tal vez no haya un límite definido que separe los tipos de autoexplicación parcial intuitivamente admisibles de los que obstaculizan la intuición, pues hasta la explicación potencial recién considerada —que es aceptable en su formulación original— podría juzgarse según bases intuitivas cuando se transforma en la versión equivalente que se acaba de dar.

El punto ilustrado por el último ejemplo se expone de manera más explícita en el teorema siguiente, que formulamos aquí sin demostración: *(6.7) Teorema.* Sea (T, C) un explanans potencial de la oración singular E. Existen entonces tres oraciones singulares, E_1, E_2 y C_1, en L tales que E equivale a la conjunción $E_1 \cdot E_2$; C equivale a la conjunción $C_1 \cdot E_1$ y E_2 puede derivarse en L de T sola.[35]

En términos más intuitivos, esto significa que si representamos la estructura deductiva de una explicación potencial dada mediante el esquema $[T,C] \rightarrow E$, entonces este esquema puede reexpresarse en la forma $[T, C_1 \cdot E_1] \rightarrow E_1 \cdot E_2$, donde E_2 se deduce de T sola, de modo que C_1 es totalmente innecesaria como premisa; en consecuencia, el esquema deductivo que estamos considerando puede reducirse a $[T, E_1] \rightarrow E_1 \cdot E_2$, que puede descomponerse en los dos esquemas deductivos $[T] \rightarrow E_2$ y $[E_1] \rightarrow E_1$. El primero de estos esquemas podría denominarse una explicación teórica pura de E_2 por T, y el segundo, una autoexplicación completa de E_1. En otras palabras: el teorema (6.7) demuestra que toda explicación cuyo explanandum es una oración singular puede descomponerse en una explicación puramente teórica y en una autoexplicación completa. Y cualquier explicación de esta clase en que el componente singular del explanans no sea completamente innecesario implica una autoexplicación parcial del explanandum.[36]

[35] En la formulación del teorema dado y en el texto siguiente, los símbolos de enunciados conectivos se emplean no sólo como signos en L sino también de manera autónoma al hablar acerca de expresiones compuestas de L. Así, cuando "S" y "T" son nombres o variables nominales de oraciones en L, su conjunción y disyunción se designarán por "$S \cdot T$" y "$S \vee T$", respectivamente; el condicional cuyo antecedente es S y el consecuente T serían designados por "$S \supset T$", y la negación de S por "$\sim S$". (A propósito, esta convención ya ha sido utilizada una vez, de modo tácito, en la nota 33.)

[36] La característica a que nos hemos referido como autoexplicación parcial debe distinguirse de lo que a veces se denomina la circularidad de la explicación científica. Esta frase ha sido empleada para expresar dos ideas distintas: a) una de ellas es la aseveración de que los principios explicativos aducidos al explicar un fenómeno específico se infieren del mismo fenómeno, de modo que todo el proceso

Prohibir la autoexplicación parcial significaría por lo tanto, la limitación de la explicación a una explicación teórica pura. Esta medida parece una restricción excesivamente severa. Por otra parte, no parece que esté justificado el intento de delimitar el grado admisible de autoexplicación mediante alguna regla especial, porque, como hemos visto, el uso corriente no proporciona ninguna orientación para tal delimitación, y porque al parecer no se obtiene ninuna ventaja sistemática mediante el trazado de una línea divisoria arbitraria. Por estas razones, nos abstenemos de introducir estipulaciones que prohiben la autoexplicación parcial.

Las condiciones establecidas en (6.5) fracasan en excluir otro tipo inaceptable de argumento explicativo que está estrechamente vinculado con la autoexplicación completa, y que deberá proscribirse por medio de una estipulación adicional. En suma, el punto es que si aceptáramos (6.5) como una definición de la explicación potencial, en lugar de considerarla como una enunciación de las condiciones necesarias, entonces, y como consecuencia de (6.6), cualquier hecho concreto dado podría explicarse por medio de una oración legal verdadera, cualquiera que fuese. Expresado de manera más explícita: si E es una oración verdadera —"La cima del monte Everest está cubierta de nieve", por ejemplo— y T es una ley —como "Todos los metales son buenos conductores del calor"—, entonces siempre existirá una oración singular verdadera C tal que E es deducible de T y C, pero no de C sola; en otras palabras, que (6.5) sea satisfecho. En efecto, sea T_s un caso particular de T elegido arbitrariamente, tal como "Si la torre Eiffel es metálica será buena conductora del calor".. Ahora bien, puesto que E es verdadera, también lo será el condicional $T_s \supset E$, y si hacemos que éste sea la oración C, entonces T,C y E satisfarán las condiciones establecidas en (6.5).

A fin de aislar las características distintivas de este tipo engañoso de explicación, examinemos un caso especialmente sencillo de este tipo objetable. Sea T_1 = "$(x)P(x)$" y E_1 = "$R(a,b)$", entonces la oración C_1 = "$P(a) \supset R(a,b)$" está formada según las instrucciones precedentes y T_1, C_1 y E_1 satisfacen las condiciones de (6.5). Sin embargo, como lo hemos visto con el ejemplo anterior, no diríamos que (T_1, C_1) constituye un explanans potencial de E_1. El fundamento teórico del veredicto puede exponerse como sigue: si la teoría T_1 sobre la que se apoya la explicación es realmente verdadera, entonces la oración C_1, que también puede ponerse en la forma "$\sim P(a) \vee R(a,b)$", solamente puede verificarse —esto es, demostrarse que es verdadera— por medio de "$R(a,b)$", es decir, E_1. En este sentido más amplio, E_1 resulta aquí explicada por sí misma. Y, por cierto, la peculiaridad que se señala priva claramente a la explicación

explicativo es circular. Esta creencia es falsa, ya que las leyes generales no pueden inferirse de oraciones singulares; b) también se ha argüido que en una explicación seria el contenido del explanandum está contenido en el del explanans. Esto es correcto, pues el explanandum es una consecuencia lógica del explanans; pero esta peculiaridad no hace trivialmente circular a la explicación científica, puesto que las leyes generales que ocurren en el explanans van mucho más allá del contenido del explanandum específico. Para un análisis más completo de la objeción sobre la circularidad, véase Feigl (1945, págs. 286 y sigs.) en donde se trata con mucha claridad este punto.

potencial propuesta para E_1 del alcance predictivo que, como señalamos en la parte I, es esencial para la explicación científica: no es posible predecir E_1 sobre la base de T_1 y C_1, pues la verdad de C_1 no puede investigarse de otra manera que no incluya la verificación de E_1. Por lo tanto, sería preciso complementar a (6.5) por la estipulación de que si (T,C) puede ser un explanans potencial de E, entonces el supuesto de que T es verdadero no debe implicar que la verificación de C requiera la verificación de E.[37]

¿Cómo podría expresarse esta idea de modo más preciso? El estudio de un ejemplo nos sugerirá una definición de la verificación para las oraciones moleculares. La oración M = "$[\sim P(a) \cdot Q(a)]$ v $R(a, b)$" puede verificarse de dos maneras distintas, sea determinando la verdad de las dos oraciones que la componen "$\sim P(a)$" y "$Q(a)$", que tienen en conjunto como consecuencia a M, o estableciendo la verdad de la oración "$R(a,b)$", que también tiene a M como consecuencia. Digamos que E es una oración básica en L si E es una oración atómica o la negación de una oración atómica en L. Puede entonces definirse de modo general la verificación de una oración molecular E diciendo que consiste en establecer la verdad de alguna clase de oraciones básicas que tenga como consecuencia a E. Por consiguiente, puede reformularse la estipulación adicional que se procura de la siguiente manera: el supuesto de que T es verdadero no debe implicar que toda clase de oraciones básicas verdaderas que tenga a C como consecuencia, también tenga a E como consecuencia.

Una rápida reflexión nos demuestra que esta estipulación puede expresarse de la siguiente manera, que evita hacer referencia a la verdad: en L, T debe ser compatible por lo menos con una clase de oraciones básicas que tenga a C, pero no a E, como consecuencia. O de modo equivalente: en L debe haber por lo menos una clase de oraciones básicas que tengan como consecuencia a C, pero no a $\sim T$ ni a E.

Si se cumple este requisito, entonces con seguridad E no puede ser consecuencia de C, porque de otro modo no podría haber ninguna clase de oraciones básicas que tenga como consecuencia a C pero no a E; de donde añadir a (6.5) la nueva condición hace superflua la segunda estipulación de (6.5) (2). Definimos ahora la explicación potencial del modo siguiente:

(6.8) Un par ordenado de oraciones, (T,C) constituye una explanans potencial de una oración singular E si, y sólo si, se cumplen las condiciones siguientes:

1) T es esencialmente generalizada y C es singular.
2) E puede derivarse de T y C juntas, en L.
3) T es compatible por lo menos con una clase de oraciones básicas que tiene como consecuencia a C y no a E.

[37] Es importante distinguir claramente entre los dos casos siguientes: a) si T es verdadera, entonces C no puede ser verdadera sin que E sea verdadera, y b) si T es verdadera, C no puede verificarse sin que E sea también verificada. La condición a) debe satisfacerse por toda explicación potencial; la condición b) mucho más restrictiva, no debe cumplirse si (T, C) ha de ser un explanans potencial de E.

La definición del concepto de explanans mediante el concepto de explanans potencial, como se formuló en (6.6.) permanece inalterada.

En función de nuestro concepto de explanans podemos dar la siguiente interpretación a la frase empleada con tanta frecuencia "este hecho es explicable por medio de esa teoría".

(6.9) Una oración singular *E* se puede explicar por medio de una teoría *T* si existe una oración singular *C* tal, que (*T*,*C*) constituya un explanans para *E*.

El concepto de explicación causal que hemos examinado aquí entraña varias generalizaciones. Una de ellas consiste en permitir que *T* incluya leyes estadísticas, pero esto requiere el fortalecimiento previo de los medios de expresión disponibles en *L*, o el uso de un aparato teórico complejo en el metalenguaje. Por otra parte, e independientemente de leyes estadísticas entre los principios explicativos, podemos reemplazar el requisito estrictamente deductivo de que *E* tiene que ser consecuencia de *T* y *C* juntas, por el inductivo más amplio de que *E* debe tener un alto grado de confirmación relacionada con la conjunción de *T* y *C*. Ambas extensiones del concepto de explicación abren importantes perspectivas y plantean una variedad de problemas nuevos. Sin embargo, en este ensayo no nos ocuparemos más de esas cuestiones.

IV. PODER SISTEMATICO DE UNA TEORIA

7. *Explicación del concepto de poder sistemático*

Las leyes y teorías científicas tienen la función de establecer conexiones sistemáticas entre los datos de nuestra experiencia, de modo que sea posible hacer la deducción de algunos de esos datos a partir de otros. La deducción se llamará explicación o predicción, según que en el momento de realizarla se sepa, o no, si los datos deducidos han ocurrido. Ahora bien, parece posible algunas veces comparar diferentes teorías, al menos de un modo intuitivo, relacionadas con su poder explicativo o predictivo: algunas teorías parecen ser poderosas en el sentido de permitir la deducción de muchos datos desde una cantidad pequeña de información inicial; otras parecen tener menos poder, pues requieren una cantidad mayor de datos iniciales o dan resultados más pobres en comparación. ¿Es posible dar una interpretación precisa a comparaciones de esta naturaleza mediante la definición, de manera completamente general, de una medida numérica del poder explicativo o predictivo de una teoría? En esta sección desarrollaremos una definición de esa clase y examinaremos algunas de sus implicaciones; en la sección siguiente, la definición será ampliada y se esbozará una teoría general del concepto en consideración.

Puesto que la explicación y la predicción tienen la misma estructura lógica, o mejor dicho, la de una sistematización deductiva, emplearemos el término neutro "poder sistemático" para referirnos al concepto propuesto. Como lo sugiere la caracterización intuitiva anterior, el poder sistemático

de una teoría T se reflejará en la relación entre la cantidad de información deducible por medio de T y la cantidad de información necesaria para esa deducción. Esta razón dependerá, evidentemente, del conjunto particular de datos o de la información a que en el caso se aplique T, y por lo tanto relativizaremos nuestro concepto en conformidad. Nuestro propósito, entonces, es construir una definición de $s(T,K)$, el poder sistemático de una teoría T con respecto a una clase finita K de datos, o el grado en que T sistematiza deductivamente la información contenida en K.

Otra vez construiremos nuestros conceptos con referencia específica al lenguaje L. Diremos que cualquier oración singular en L expresa un dato potencial, y K se interpretará, en consecuencia, como una clase finita de oraciones singulares.[38] T se interpretará en un sentido mucho más amplio que en las secciones anteriores; puede ser cualquier oración en L, sea esencialmente generalizado o no. Se adopta esta convención liberal para la mejor generalidad y simplicidad de las definiciones y teoremas que se desarrollarán.

Para obtener valores entre 0 y 1 inclusive, podríamos tratar de identificar ahora $s(T,K)$ con el porcentaje de aquellas oraciones en K deducibles de las restantes por medio de T. Así, si $K_1 = \{$"$P(a)$", "$Q(a)$", "$\sim P(b)$", "$\sim Q(b)$", "$Q(c)$", "$\sim P(d)$"$\}$, y $T_1 =$ "$(x)[P(x) \supset Q(x)]$", entonces exactamente la segunda y tercera oraciones en K_1 son deducibles por medio de T_1 a partir de las restantes, y más precisamente de la primera y cuarta. En consecuencia, podríamos considerar en formular $s(T_1,K_1) = 2/6 = 1/3$. Pero entonces, para la clase $K_2 = [$"$P(a) \cdot Q(a)$", "$\sim P(b) \cdot \sim Q(b)$", "$Q(c)$", "$\sim P(d)$"$]$, la misma T_1 tendría 0 como valor s, aunque K_2 contuviera exactamente la misma información que K_1; del mismo modo, para otra formulación más de la misma información, como $K_3 = [$"$P(a) \cdot \sim Q(b)$", "$Q(a) \cdot \sim P(b)$", "$Q(c)$", "$\sim P(d)$"$]$, T_1 tendría un valor s igual a 1/4, y así sucesivamente. Pero lo que buscamos es una medida del grado en que una teoría dada sistematiza deductivamente un conjunto dado de información fáctica, es decir, un cierto contenido al margen de la estructura y agrupación particulares de las oraciones en que ese contenido esté expresado. Emplearemos, por lo tanto, un método que representa el contenido de toda oración singular o clase de oraciones singulares como compuesto de ciertas porciones mínimas de información determinadas unívocamente. Si aplicamos nuestra idea general a estas porciones, obtendremos una medida del poder sistemático de T en K que es independiente de la manera en que está formulado el contenido de K. Las oraciones que expresan esas porciones mínimas de información se llamarán oraciones míni-

[38] Como lo muestra esta estipulación, aquí el término "dato" abarca tanto datos reales como potenciales. La convención de que cualquier oración singular expresa un dato potencial es plausible especialmente si los predicados primitivos de L se refieren a atributos cuya presencia o ausencia en casos específicos pueden indagarse mediante la observación directa. En este caso, se puede considerar que cada oración singular en L expresa un dato potencial, en el sentido de que describe un estado de cosas lógicamente posible, cuya existencia se puede investigar mediante la observación directa. El supuesto de que los primitivos de L expresan atributos directamente observables no es esencial, sin embargo, para la definición y la teoría formal del poder sistemático expuestas en las secciones 7 y 8.

mas, y una definición explícita de este concepto auxiliar hará posible una formulación exacta del procedimiento propuesto. Ahora desarrollaremos este punto.

Si, como supondremos aquí, el vocabulario de L contiene números fijos y finitos de constantes individuales y de constantes predicativas, entonces sólo se puede formular en L un cierto número finito —digamos n— de oraciones atómicas distintas. Por oración mínima en L entenderemos una disyunción de cualquier número de k $(O \leqq k \leqq n)$ de oraciones atómicas diferentes y las negaciones de las n-k restantes. Resulta claro que las oraciones atómicas n determinan oraciones mínimas 2^n. Así un lenguaje L_1 contiene exactamente una constante individual "a" y exactamente dos predicados primitivos "P" y "Q", ambos de grado 1; entonces L_1 contiene dos oraciones atómicas "$P(a)$" y "$Q(a)$", y cuatro oraciones mínimas: "$P(a)$ v $Q(a)$", "$P(a)$ v $\sim Q(a)$", "$\sim P(a)$ v $Q(a)$", "$\sim P(a)$v $\sim Q(a)$". Si otro lenguaje, L_2, contiene además del vocabulario de L_1, una segunda constante individual, "b", y un predicado "R" de grado 2, entonces L_2 contiene ocho oraciones atómicas y 256 oraciones mínimas tales como "$P(a)$ v $P(b)$ v $\sim$ $Q(a)$ v $Q(b)$ v $R(a, a)$ v $R(a, b)$ v $\sim$ R(b, a) v $\sim$ $R(b, b)$".

La expresión "oración mínima" indica que los enunciados estudiados son las oraciones singulares en L de menor contenido diferente de cero, lo que significa que toda oración singular en L que se deduce de una oración mínima equivale a esa oración mínima o formalmente verdadera en L. Las oraciones mínimas tienen, por cierto, consecuencias diversas de ellas mismas que no son formalmente verdaderas en L, pero éstas no son de forma singular; "$(Ex)(P(x)$ v $Q(x))$" es una consecuencia de "$P(a)$ v $Q(a)$" en el lenguaje L_1 al que nos hemos referido.

Además, no hay dos oraciones mínimas que tengan ninguna consecuencia común que no sea formal o lógicamente verdadera en L; en otras palabras, los contenidos de dos oraciones mínimas cualesquiera son recíprocamente excluyentes.

En virtud de los principios del cálculo oracional, toda oración singular que no sea formalmente verdadera en L puede transformarse en una conjunción de oraciones mínimas unívocamente determinadas; esta conjunción se denominará forma mínima normal de la oración. Así, por ejemplo, en el lenguaje L_1, al que nos referimos más arriba, las oraciones "$P(a)$" y "$Q(a)$" tienen las formas mínimas normales "$[P(a)$ v $Q(a)]$" · "$[P(a)$ v $\sim$ $Q(a)]$" y "$[P(a)$ v $Q(a)]$ · $[\sim P(a)$ v $Q(a)]$", respectivamente: en L_2, cada una de esas mismas oraciones tiene formas mínimas normales consistentes en 128 oraciones mínimas conjuntivas. Si una oración es formalmente verdadera en L, su contenido es cero, y no puede representarse por una conjunción de oraciones mínimas. Será conveniente, sin embargo, decir que la forma mínima normal de una oración formalmente verdadera en L es la conjunción vacua de oraciones mínimas, que no contiene ningún término.

Como consecuencia del principio mencionado, cualquier clase de oraciones singulares que no sean todas formalmente verdaderas pueden representarse por una oración de forma mínima normal. Podemos expresar

ahora la idea básica ya esbozada para la explicación del concepto de poder sistemático mediante la siguiente definición:

(7.1) Sea T cualquier oración en L y K cualquier clase finita de oraciones singulares en L que no son todas formalmente verdaderas. Si K' es la clase de oraciones mínimas que ocurren en la forma mínima normal de K, consideremos todas las divisiones de K' dentro de dos subclases recíprocamente excluyentes, K'_1 y K'_2, de modo que toda oración K'_2 sea deducible de K'_1 por medio de T. Cada división de este tipo determina una relación n (K'_2) / n (K'), es decir, la cantidad de oraciones mínimas en K'_2 dividida por el número total de oraciones mínimas en K'. Entre los valores de estas relaciones, debe haber uno que sea mayor; $s(T, K)$ debe ser igual a esa relación máxima. (Nótese que si todos los elementos de K fueran formalmente verdaderos, $n(K')$ sería 0 y la razón a que nos referimos no se podría definir.)

Ejemplo: Supongamos que L_1 contiene una sola constante individual, "a", y sólo dos predicados, "P" y "Q", ambos de grado 1. En L_1 hagamos $T =$ "$(x)[P(x) \supset Q(x)]$", $K =$ ["$P(a)$", "$Q(a)$"]. Entonces tenemos $K' =$ ["$P(a) \vee Q(a)$", "$P(a) \vee \sim Q(a)$", "$\sim P(a) \vee Q(a)$"]. A partir de la subclase K'_1, que consiste en los dos primeros elementos de K' —que juntos equivalen a "$P(a)$"— podemos deducir, por medio de T, la oración "$Q(a)$", y de ahí por pura lógica, el tercer elemento de K'; éste constituye el único elemento de K'_2. No es posible ninguna sistematización "mejor", de donde $s(T, K) = 1/3$.

Nuestra definición es independiente del problema, que deja sin resolver, de si para una K' dada no podrían existir diferentes divisiones, cada una de las cuales daría el valor máximo para $n(K'_2)/n(K')$. En realidad, esto nunca puede suceder: si existe exactamente una subdivisión óptima de K' dada. Este hecho es un corolario de un teorema general, del que nos ocuparemos ahora. Se advertirá que en el último ejemplo, K'_2 puede deducirse de T sola, sin emplear K_1 como premisa; por cierto, "$\sim P(a) \vee Q(a)$" no es más que un caso de sustitución de la oración "$(x)[\sim P(x) \vee Q(x)]$", equivalente a T. El teorema que ahora formularemos y que puede parecer sorprendente al principio, muestra que esta observación se aplica de manera análoga en todos los otros casos.

(7.2) *Teorema.* Si T es una oración cualquiera, K' una clase de oraciones mínimas, y K'_2 una subclase de K' tal que toda oración en K'_2 sea deducible por medio de T de la clase K - K'_2, entonces toda oración de K_2' es deducible de T sola.

La demostración, en síntesis, es como sigue: puesto que los contenidos de dos oraciones mínimas diferentes cualesquiera son recíprocamente excluyentes, deben serlo también los contenidos de K'_1 y K'_2 que no tienen ninguna oración mínima en común. Pero puesto que las oraciones de K'_2 se deducen de K'_1 y T juntas, deben por lo tanto deducirse de T sola.

Señalamos las siguientes consecuencias de nuestro teorema:

(7.2a) *Teorema.* En cualquier clase K' de oraciones mínimas, la ma-

yor subclase deducible de las restantes por medio de una oración T es idéntica a la clase de aquellos elementos en K' que son deducibles de T sola.

(7.2b) *Teorema.* Si T es cualquier oración, K una clase de oraciones singulares, K' la clase equivalente de oraciones mínimas, y K'_t la clase de aquellas de entre estas últimas, deducibles de T sola, entonces el concepto s definido en (7.1) satisface la siguiente ecuación:

$$s(T, K) = n(K'_t) / n(K')$$

8. *Poder sistemático y probabilidad lógica de una teoría. Generalización del concepto de poder sistemático*

El concepto de poder sistemático está estrechamente vinculado con el grado de confirmación, o probabilidad lógica, de una teoría. El estudio de esta relación esclarecerá la definición de s propuesta, sugerirá ciertas maneras de generalizarla y conducirá finalmente a una teoría general de poder sistemático que es formalmente análoga a la de la probabilidad lógica.

El concepto de probabilidad lógica, o grado de confirmación, es el concepto central de la lógica inductiva. Carnap,[39] Helmer, Hempel y Oppenheim[40] han propuesto diferentes definiciones explícitas de este concepto, para lenguajes de estructura similar a la de nuestro lenguaje modelo.

Mientras que la definición de s propuesta en la sección anterior se apoya en el concepto de oración mínima, el concepto básico en la construcción de una medida de la probabilidad lógica es el de la descripción de estado o, como también dijimos, el de la oración máxima. Una oración máxima es el dual[41] de una oración mínima en L; es una conjunción de oraciones atómicas diferentes $k (0 \leq k \leq n)$, y de las negaciones de las oraciones atómicas restantes $n - k$. En un lenguaje con oraciones atómicas n, existen 2^n descripciones de estado. Así, por ejemplo, el lenguaje L_1 repetidamente mencionado en el párrafo 7 contiene las cuatro siguientes oraciones máximas: "$P(a) \cdot Q(a)$", "$P(a) \cdot \sim Q(a)$", "$\sim P(a) \cdot Q(a)$", "$\sim P(a) \cdot \sim Q(a)$".

El término "oración máxima" es para indicar que las oraciones en cuestión son las oraciones singulares de máximo contenido no universal en L, que significa que toda oración singular en L que tiene como consecuencia una oración máxima es equivalente a esa oración máxima o formalmente falsa en L.

[39] Véase especialmente 1945, 1495a y 1947.

[40] Véase Helmer y Oppenheim (1945); Hempel y Oppenheim (1945). Ciertos aspectos generales de la relación entre la confirmación de una teoría y su éxito predictivo o sistemático se examinan en Hempel, 1945, parte II, sec. 7 y 8. La definición de s desarrollada en este volumen establece una contraparte cuantitativa de lo que en este trabajo se caracteriza, en términos no numéricos, como el criterio predictivo de la confirmación.

[41] Para una definición y análisis de este concepto véase, por ejemplo, Church (1942, pág. 172).

Como hemos visto, toda oración singular puede representarse en una forma normal conjuntiva, o mínima, es decir, como una conjunción de ciertas oraciones mínimas unívocamente determinadas; de manera similar, toda oración singular puede expresarse también en una forma normal disyuntiva o máxima, es decir, como una disyunción de ciertas oraciones máximas unívocamente determinadas. En el lenguaje L_1, por ejemplo, "$P(a)$" tiene la forma normal mínima "$[P(a) \vee Q(a)] \cdot [P(a) \vee \sim Q(a)]$" y la forma normal máxima "$[P(a) \cdot Q(a)] \vee [P(a) \cdot \sim Q(a)]$"; la oración "$P(a) \supset Q(a)$" tiene la forma normal mínima "$\sim P(a) \vee Q(a)$" y la forma normal máxima "$[P(a) \cdot Q(a)] \vee [\sim P(a) \cdot Q(a)] \vee [\sim P(a) \cdot \sim Q(a)]$"; la forma normal mínima de una oración formalmente verdadera es la conjunción vacua, mientras que su forma normal máxima es la disyunción de la totalidad de las cuatro descripciones de estado en L_1. La forma normal mínima de cualquier oración formalmente falsa es la conjunción de todas las cuatro oraciones mínimas en L_1, en tanto que su forma normal máxima es la disyunción vacua.

La forma mínima normal de una oración singular es un índice adecuado de su contenido, pues representa a la oración como una conjunción de componentes estándares cuyos contenidos son mínimos y recíprocamente excluyentes. La forma normal máxima de una oración es un índice adecuado de su amplitud, es decir, expresado de un modo intuitivo, de la variedad de sus diferentes realizaciones posibles, o de la variedad de aquellos estados posibles del mundo que, si se realizaran, harían verdadero ese enunciado. En realidad, cada oración máxima se puede considerar como la descripción, tan completa como sea posible en L, de un probable estado del mundo; y las descripciones de estado que constituyen la forma normal máxima de una oración singular dada no hace sino registrar aquellos estados, entre todos los posibles, que harían verdadera la oración.

Del mismo modo que los contenidos de dos diferentes oraciones mínimas cualesquiera, las amplitudes de dos oraciones máximas cualesquiera son recíprocamente excluyentes: ningún estado posible del mundo puede hacer verdaderas dos oraciones máximas diferentes, porque dos oraciones máximas cualesquiera son obviamente incompatibles entre sí.[42]

La amplitud y el contenido de una oración varían en proporción inversa. Cuanto más expresa una oración, menor es la variedad de sus posibles realizaciones, y a la inversa. Esta relación se refleja en el hecho de que cuanto mayor sea el número de componentes de la forma normal mínima de una oración singular, menor es el número de componentes de su forma normal máxima, y a la inversa. En efecto, si la forma normal mínima de una oración singular U contiene m_U de $m = 2^n$ oraciones mínimas en L, entonces su forma normal máxima contiene $l_U = m - m_U$ de las oraciones máximas m en L. Esto queda ilustrado por nuestros cuatro últimos ejemplos, donde $m = 4$, y $m_U = 2, 1, 0, 4$, respectivamente.

Las observaciones anteriores sugieren que el contenido de toda oración singular U puede medirse mediante el correspondiente número m_U o por

42 Se puede encontrar un análisis más detallado del concepto de amplitud en Carnap (1945, sec. 2 y en 1942, sec. 18 y 19) donde se examina extensamente la relación entre amplitud y contenido.

medio de alguna magnitud proporcional a él. Ahora bien, será conveniente restringir los valores de la función que expresa la medida del contenido al intervalo de 0 a 1 inclusive; y por lo tanto, definimos una medida, $g_1(U)$ por el contenido de cualquier oración singular en L por medio de la fórmula

(8.1) $$g_1(U) = m_U/m$$

A cualquier clase finita K de oraciones singulares le asignamos, como medida $g_1(K)$ de su contenido, el valor $g_1(S)$, donde S es la conjunción de los elementos de K.

En virtud de esta definición podemos volver a escribir la ecuación del teorema (7.2b):

$$s(T, K) = g_1(K'_t)/g_1(K')$$

Aquí, K'_t es la clase de todas aquellas oraciones mínimas de K' que son consecuencias de T. En el caso especial en que T sea una oración singular, K'_t, será, en consecuencia, equivalente a $T \text{v} S$, donde S es la conjunción de todos los elementos de K'. Por lo tanto, la ecuación anterior puede transformarse en

(8.2) $$s(T, S) = g_1(T \text{ v } S)/g_1(S)$$

Esta fórmula es válida cuando T y S son oraciones singulares, y S no es formalmente verdadera. Tiene una semejanza sorprendente con el esquema general para definir la probabilidad lógica de T con respecto a S:

(8.3) $$p(T, S) = r(T \cdot S)/r(S)$$

Aquí $r(U)$ es, para cualquier oración U en L, una medida de la amplitud de U; T es cualquier oración en L, y S cualquier oración en L con $r(S) \neq 0$.

Las diversas definiciones específicas del concepto de probabilidad lógica que se han propuesto concuerdan esencialmente con la pauta general exhibida en (8.3),[43] pero difieren en la selección de una función de medida específica para las amplitudes, es decir, difieren en la definición de r. Una idea que surge por sí misma es la de asignar, a cualquier oración singular U cuya forma máxima normal contiene oraciones máximas l_U, la medida de la amplitud:

(8.4) $$r_1(U) = l_U/m$$

definida, evidentemente, en estricta analogía a la medida del contenido g_1 aplicable a oraciones singulares, que se ha introducido en (8.1). Para toda oración singular U la suma de las dos medidas es igual a la unidad:

(8.5) $$r_1(U) + g_1(U) = (l_U + m_U)/m = 1$$

Como lo ha demostrado Carnap, sin embargo, la medida de la ampli-

[43] En la teoría de la probabilidad lógica de Carnap, $p(T, S)$ se define en ciertos casos como el límite que asume la función $r(T \cdot S) / r(S)$ en condiciones especificadas (véase Carnap, 1945, pág. 75); pero nosotros nos abstendremos aquí de considerar esta generalización del tipo de definición representado por (8.3).

tud r_1 confiere al correspondiente concepto de probabilidad lógica, es decir, al concepto p_1 definido mediante ella según el esquema (8.3), ciertas características que son incompatibles con el significado de la probabilidad lógica intentado;[44] tanto Carnap como Helmer y los autores de este capítulo han sugerido ciertas funciones de medida de amplitud que pueden servir de alternativa, y que conducen a conceptos de probabilidad y de grado de confirmación más satisfactorios. Aunque no es necesario entrar en detalles, las observaciones generales que mencionamos a continuación parecen indicadas para preparar el análisis subsiguiente.

La función r_1 mide la amplitud de una oración singular esencialmente mediante el recuento del número de oraciones máximas en su forma normal máxima: así, concede igual peso a todas las oraciones máximas (la definición (8.1) trata de manera análoga las oraciones mínimas). Las definiciones alternativas a las que nos referimos se basan en un procedimiento diferente. Carnap, en particular, formula la regla que asigna un peso específico, es decir, un valor específico r a cada oración máxima, pero estos pesos no son iguales para todas las oraciones máximas. Luego define la medida de la amplitud de cualquier otra oración singular como la suma de las medidas de las oraciones máximas que la componen. Según la función así obtenida —llamémosla r_2— Carnap define el correspondiente concepto de probabilidad lógica, que llamaremos p_2, para oraciones singulares T, S de acuerdo con el esquema (8.3) : $p_2(T, S) = r_2(T \cdot S)/r_2(S)$. Las definiciones de r_2 y p_2 se extienden luego, mediante ciertos procesos limitadores, a los casos en donde T y S ya no son ambas singulares.[45]

Se puede advertir fácilmente que así como la función r_1 definida en (8.5) no es sino una de entre una infinidad de posibles medidas de la amplitud, del mismo modo la función análoga g_1 definida en (8.1) no es sino una de entre una infinidad de posibles medidas del contenido; y así como cada medida de la amplitud puede servir para definir, según el esquema (8.3), una medida de la probabilidad lógica correspondiente, del mismo modo cada función de medida del contenido puede servir para definir, por medio del esquema ilustrado en (8.2), una medida correspondiente del poder sistemático. Un método que se sugiere para obtener otras

[44] 1945, págs. 80-81.

[45] El enfoque alternativo sugerido por Helmer y los autores de este capítulo implica el uso de una función de medida de amplitud r_1 que depende, de manera especificada, de la información r disponible; en consecuencia, la medida de la amplitud de cualquier oración U está determinada sólo si se da una oración I que exprese la información empírica disponible. De acuerdo con esta función de medida de la amplitud, el concepto de grado de confirmación, *dc*, puede definirse por medio de una fórmula similar a (8.3). El valor de $dc(T, S)$ no se define, sin embargo, en ciertos casos en donde S está generalizado, como lo ha señalado McKinsey (1946); además, el concepto *dc* no satisface todos los teoremas de la teoría elemental de la probabilidad (véase el análisis de este punto en los dos primeros artículos mencionados en la nota [40]); por lo tanto, el grado de confirmación de una teoría con respecto a una comprobación dada no es una probabilidad en el sentido estricto de la palabra. Por otra parte, la definición de *dc* a que nos referimos aquí tiene ciertos rasgos metodológicamente deseables y, en consecuencia, podrían ser interesante para interpretar un concepto vinculado al poder sistemático mediante la función r_1 de medida de la amplitud. En este capítulo, sin embargo, no nos ocuparemos de este problema.

funciones de medida del contenido consiste en seleccionar alguna medida de amplitud r distinta de r_1 y luego *definir* una medida de contenido correspondiente g, en función de aquélla, mediante la fórmula

$$(8.6) \qquad g(U) = 1 - r(U)$$

de modo que g y r satisfagan el análogo de (8.5) por definición. La función g así definida conducirá a su vez, por vía de una definición análoga a (8.2), a un concepto g correspondiente. Consideremos este procedimiento con más detalle.

Supongamos que se da una función r que satisfaga los requisitos habituales de las medidas de amplitud:

(8.7)
1. $r(U)$ está unívocamente determinada para todas las oraciones U en L.
2. $0 \leqq r(U) \leqq 1$ para toda oración U en L.
3. $r(U) = 1$ si la oración U es formalmente verdadera en L y de esa manera tiene amplitud universal.
4. $r(U_1 \vee U_2) = r(U_1) + r(U_2)$ para dos oraciones cualesquiera U_1 y U_2 cuyas amplitudes son recíprocamente excluyentes, es decir, cuya conjunción es formalmente falsa.

En función de la medida de amplitud dada, definamos la correspondiente medida de contenido g por medio de (8.6). Entonces se puede demostrar que g satisface las siguientes condiciones:

(8.8)
1. $g(U)$ está unívocamente determinada para todas las oraciones U en L.
2. $0 \leqq g(U) \leqq 1$ para toda oración U en L.
3. $g(U) = 1$ si la oración U es formalmente falsa en L y tiene así contenido universal.
4. $g(U_1 \cdot U_2) = g(U_1) + g(U_2)$ para dos oraciones cualesquiera U_1 y U_2 cuyos contenidos son recíprocamente excluyentes, es decir, cuya disyunción es formalmente verdadera.

En analogía con (8.2) podemos definir a continuación, por medio de g, una función s correspondiente.

$$(8.9) \qquad s(T, S) = g(T \vee S)/g(S)$$

Esta función está determinada por toda oración T y por toda oración S siendo $g(S) \neq 0$, mientras que la definición de poder sistemático dada en el parágrafo 7 estaba limitada a aquellos casos en que S es singular y no formalmente verdadera. Por último, nuestra medida de amplitud r determina una función de probabilidad correspondiente, en virtud de la definición

$$(8.10) \qquad p(T, S) = r(T \cdot S)/r(S)$$

Esta fórmula determina la función p para cualquier oración T y para toda oración S, siendo $r(S) \neq 0$.

De esta manera, toda medida de amplitud r que satisfaga (8.7) determina unívocamente una medida de contenido g correspondiente, que satisface (8.8), una función s correspondiente, definida por (8.9), y una función p correspondiente, definida por (8.10). Como consecuencia de (8.7) y de (8.10), se puede demostrar que la función p satisface las leyes elementales de la teoría de la probabilidad, especialmente las que se mencionan en (8.12), a continuación; y en virtud de ellas, es posible establecer una relación muy simple que se obtiene, para cualquier medida dada de amplitud r, entre los conceptos $p(T, S)$ y $s(T,S)$ correspondientes. Tenemos

(8.11)
$$\begin{aligned} s(T, S) &= g(T\text{v}S)/g(S) \\ &= [1 - r(T\text{v}S)]/[1 - r(S)] \\ &= r[\sim (T\text{v}S)]/r(\sim S) \\ &= r(\sim T \cdot \sim S)/r(\sim S) \\ &= p(\sim T, \sim S) \end{aligned}$$

Pasamos a enunciar, sin demostrarlos, algunos teoremas que aluden a p y s, y que son consecuencia de nuestros supuestos y definiciones; son válidos en todos los casos en que los valores de p y s que se mencionan existen, es decir, donde el valor-r del segundo argumento de p, y el valor-g de los argumentos segundos de s, no es O.

(8.12)

1) a. $0 \leqq p(T, S) \leqq 1$

b. $0 \leqq s(T, S) \leqq 1$

2) a. $p(\sim T,S) = 1 - p(T, S)$

b. $s(\sim T, S) = 1 - s(T, S)$

3) a. $p(T_1 \text{ v } T_2, S) = p(T_1, S) + p(T_2, S) - p(T_1 \cdot T_2, S)$

b. $s(T_1 \cdot T_2, S) = s(T_1, S) + s(T_2, S) - s(T_1 \text{ v } T_2, S)$

4) a. $p(T_1 \cdot T_2, S) = p(T_1, S) \cdot p(T_2, T_1 \cdot S)$

b. $s(T_1, \text{v } T_2, S) = s(T_1, S) \cdot s(T_2, T_1 \text{ v } S)$

En este agrupamiento, estos teoremas ilustran la relación de correspondencia dual que se obtiene entre p y s. Una caracterización general de esta correspondencia se da en el siguiente teorema, que puede comprobarse sobre la base de (8.11), y que se enuncia aquí de una manera ligeramente informal con el fin de evitar el fastidio de las formulaciones extensas.

(8.13) *Teorema del dualismo.* A partir de cualquier fórmula general demostrable que exprese una igualdad o una desigualdad con respecto a p se obtiene una fórmula demostrable respecto de s, si se reemplaza "p" en todos los casos por "s" y "·" y "v" se sustituyen recíprocamente. La misma sustitución, y el reemplazo de "s" por "p" recíprocamente, transforma todo teorema que exprese una igualdad o una desigualdad respecto de s en un teorema referente a p.

Comenzamos nuestro análisis del poder sistemático de una teoría con relación a una clase de datos mediante la interpretación de este concepto, en el parágrafo 7, como una medida de la relación óptima de datos entre

aquellos que son deducibles de los restantes por medio de esa teoría. La elaboración sistemática de esta idea ha conducido a la definición, en esta sección, de un concepto más general de poder sistemático, que resultó ser la contraparte dual del concepto de probabilidad lógica. Esta extensión de nuestra interpretación original da como resultado una teoría más sencilla y más inclusiva que la que se hubiera podido lograr sobre la base de nuestra definición inicial.

Pero la teoría del poder sistemático, tanto en su expresión más estrecha como en la versión generalizada es, lo mismo que la teoría de la probabilidad lógica, de carácter puramente formal, y para una aplicación significativa de cualquiera de estas dos teorías en la epistemología o en la metodología de la ciencia se requiere la solución de ciertos problemas fundamentales que conciernen a la estructura lógica del lenguaje de la ciencia y la interpretación de sus conceptos. Una necesidad urgente es la elucidación más amplia del requisito de primitivos cualitativos puros en el lenguaje de la ciencia; otro problema crucial es elegir, entre una infinidad de posibilidades formales, una medida adecuada de la amplitud *r*. La complejidad y la dificultad de los problemas que surgen en estos contextos se han manifestado en investigaciones referentes al tema;[46] sólo se puede esperar que a los progresos alcanzados de la teoría formal seguirá pronto la solución de esos problemas aún pendientes, lo que esclarecerá las condiciones de una aplicación seria de las teorías de la probabilidad lógica y del poder sistemático.

POSTSCRIPTUM (1964) A LA LOGICA DE LA EXPLICACION

El ensayo precedente ha sido ampliamente comentado en la bibliografía filosófica. Gran parte de la discusión se centra sobre la concepción general expuesta en la Parte I acerca de la explicación por medio de la subsunción deductiva a leyes o principios teóricos. Algunos comentaristas hasta parecen atribuirme la opinión de que todas las explicaciones científicas adecuadas deben ser de este tipo, a pesar del hecho de que en los párrafos finales de las secciones 2 a 6 del ensayo, como también en la sección 5.3 del capítulo titulado "La función de las leyes generales en historia" se da por sentado otro tipo de explicación que invoca las leyes estadísticas y de probabilidades. Sin embargo, la lógica de esa explicación no se explora en profundidad en ninguno de esos dos trabajos. Un intento para llenar este vacío lo constituye la sección 3 del capítulo "Aspectos de la explicación científica" con que se cierra el presente volumen. Ese ensayo también incorpora mis respuestas a algunos de los estimulantes comentarios y críticas que se han dirigido a los dos estudios anteriores.

En esta postdata, me limitaré a enumerar algunas omisiones en las ideas desarolladas en la Parte III del ensayo precedente.

(1) Como lo señalara correctamente E. Nagel,[47] la definición (6.3b) del concepto de ley derivativa es demasiado restricta ya que, contraria-

[46] Véase especialmente Goodman (1946, 1947, 1947a) y Carnap (1947).
[47] E. Nagel: *The Structure of Science*. Nueva York, 1961, pág. 58.

mente a la intención indicada en la sección 5, excluye leyes tales como las de Galileo y Kepler de la condición de leyes derivativas. Esto se debe a que aquellas generalizaciones no pueden derivarse sólo de las leyes fundamentales de la mecánica y de la gravitación de Newton, lo cual, en efecto, podría hacerse meramente sustituyendo los términos constantes por variables que ocurren en la última. En realidad, la derivación requiere premisas adicionales que no tienen el carácter de leyes fundamentales. En el caso de la ley de Galileo, por ejemplo, ellas incluyen enunciados que especifican la masa y el radio de la Tierra. (En verdad, aun con la ayuda de premisas adicionales, las leyes de Galileo y de Kepler no pueden derivarse estrictamente de los principios newtonianos: sólo son aproximaciones de enunciados que son así derivables. Sin embargo este punto, que se profundiza en la sección 2 de "Aspectos de la explicación científica", evidentemente no disminuye la fuerza del argumento de Nagel.)

Nagel también señala que si la definición (6.3b) se modificara como para favorecer el uso de otras premisas adicionales que no tengan carácter de ley, entonces ciertos ejemplos inapropiados podrían postularse como leyes derivativas. De hecho esto se aplicaría por ejemplo a la frase "Cada manzana que se encuentra ahora en esta canasta, es roja" que puede deducirse de la ley (putativa) "Todas las manzanas Winesap son rojas" en conjunción con la premisa "Cada manzana que se encuentra ahora en esta canasta es de la variedad Winesap". Nagel ilustra este punto con la frase "Todas las tuercas del automóvil de Smith están oxidadas" que puede deducirse de la ley "El hierro expuesto al oxígeno se oxida" en conjunción con premisas particulares adecuadas.

Lo que excluye a generalizaciones como las dos recién mencionadas de la categoría de leyes potenciales praecería ser su alcance limitado: cada una pertenece sólo a un número finito de objetos. Esta observación sugiere que el requisito de no limitar el alcance (que en la sección 6 se impone a las frases fundamentales con carácter de ley) debería extenderse también a las frases derivativas con carácter de ley. En efecto, Nagel exige que las frases con carácter de ley en general sean "universales no restringidas", es decir, que su "poder de predicación" no caiga dentro de "una región espacial fija o un período determinado".[48] Con esta formulación del requisito propuesto, empero, podría suceder que una frase dada es descalificada mientras que otra, lógicamente equivalente, no lo es. Por ejemplo, los dos universales restringidos que consideramos recién son lógicamente equivalentes a los siguientes alcances de la predicción, que evidentemente satisfacen la condición propuesta por Nagel: "Cualquier cosa que no sea roja no es una manzana en esta canasta" y "Cualquier objeto que no sea oxidable no es un tuerca en el automóvil de Smith".

Esta dificultad se evita si al requerimiento del alcance se le da la siguiente forma. Exceptuando las verdades puramente lógicas (que son equivalentes de "*Pa* v ∼ *Pa*"), los enunciados con carácter de ley no deben tener un alcance finito en el sentido de ser lógicamente equivalentes a alguna conjunción finita de frases aisladas sobre casos particulares (como en "la

[48] Obra citada, pág. 59.

manzana *a* es roja y la manzana *b* es roja y la manzana *c* es roja"). O bien, para decirlo con mayor precisión y sucintamente: en esencia, deben ser universales. Es evidente que si una frase satisface esta condición, también lo hará cualquier equivalente lógico de la misma.

Esta condición, que las definiciones (6.3a) y (6.3) realmente imponen a las frases fundamentales y derivativas con carácter de ley, es tratada más a fondo en la sección 2.1 de "Aspectos de la explicación científica". Pero si bien es una condición necesaria para las frases con carácter de ley, es demasiado débil para evitar en su totalidad la dificultad señalada por Nagel. De hecho, no descarta las dos generalizaciones indeseables recién consideradas: ninguna de ellas puede transformarse de manera equivalente en una conjunción finita de frases singulares sobre determinadas manzanas o tuercas, ya que las frases ni siquiera indican cuántas manzanas hay en la canasta o cuántas tuercas en el automóvil de Smith. Tampoco proporciona una lista de denominaciones para cada uno de los objetos nombrados, como lo exigiría la transformación. De ahí que sería un desiderátum importante encontrar una versión satisfactoria de la condición del alcance que exige más bien una frase con carácter de ley que la de ser esencialmente universal.

(2) Torno ahora a las deficiencias de la definición, (6.8) de un explanans potencial. Esa definición, como ya lo advirtiera hace unos cuantos años, es demasiado incluyente ya que, en cierto sentido que pasaré a ejemplificar, favorece por sí sola la explicación de cualquier hecho y posibilita la generación de una teoría potencialmente explicativa para cualquier hecho dado a partir de cualquier frase esencialmente generalizada. Considérese, por ejemplo, el argumento

$$\text{(2a)}\qquad \frac{(x)\,Px \qquad Qa}{Qa} \qquad \text{o bien, resumiendo} \qquad \frac{T \qquad C}{C}$$

Toma la forma de ser completamente autoexplicativa y por lo tanto queda descartada por la condición (3) en la definición (6.8). Pero su explanans puede volver a formularse de un modo equivalente en la forma aceptada en (6.8), que arroja el siguiente argumento:

$$\text{(2b)}\qquad \frac{(x)\,(Px\cdot Qa) \qquad Qa \vee \sim Qa}{Qa} \qquad \text{o bien, resumiendo} \qquad \frac{T' \qquad C'}{C}$$

Este argumento claramente satisface las condiciones (1) y (2) en (7.8). Pero también satisface la condición (3); dado que T' es compatible con la clase que contiene la frase básica "*Pb*" como único elemento; que además esa clase tiene a C' y no a C como consecuencia lógica.

Este defecto puede eliminarse limitando T en la definición (6.8) a frases puramente generalizadas. Sin embargo, ésta es una restricción poco conveniente ya que la definición también estaba destinada a abarcar la explicación por medio de leyes y teorías derivativas.

(3) Aun si estuviéramos dispuestos a pagar ese precio, la versión modificada de (6.8) igual arrojaría consecuencias bastante inaceptables. Un

agudo estudio crítico de Eberle, Kaplan y Montague[49] destaca este punto y demuestra que virtualmente cualquier teoría fundamental provee una explicación en el sentido de (6.8) para virtualmente cualquier hecho. Los autores lo establecen por medio de cinco teoremas, cada uno de los cuales presenta tales relaciones de explicabilidad para alguna gran clase de casos en los cuales normalmente la teoría se consideraría irrelevante para el hecho que debe ser explicado.

El primero de esos teoremas, por ejemplo, es el siguiente: Sea T una ley fundamental y E un frase singular verdadera, ninguna de las cuales puede ser probada desde un punto de vista lógico en el idioma L. Además: que las dos frases no tengan ningún predicado común, de modo que, hablando intuitivamente, T opera con un tema totalmente diferente de E. Entonces, dada solamente la disponibilidad de un abastecimiento adecuado de otras constantes y predicados en L, existe una ley fundamental T' que es lógicamente derivable de T y por la cual E es explicable en el sentido de la definición (6.9). Por ejemplo, asumamos que T es "$(x)\,Fx$" y que E es "Ha"; considérese entonces la frase

$$T' : (x)\,(y)\,[Fx \vee (Gy \supset Hy)]$$

Su forma es puramente universal y es derivable de T; entonces, es verdadera, dado que por hipótesis T es una ley y por lo tanto verdadera. En consecuencia T' es una ley fundamental. Considérese ahora la frase

$$C\colon (Fb \vee \sim Ga) \supset Ha$$

Esta frase es singular y consecuencia de E; por lo tanto es verdadera dado que por hipótesis E es verdadera. Como podrá fácilmente verificarse ahora (T',C) forma un explanans potencial (que es verdadero) para E en el sentido de (6.8).

Es una satisfacción para mí poder decir como conclusión que es posible modificar las definiciones (6.8) y (6.9) para prevenir estas consecuencias inhabilitantes. Uno de los métodos ha sido señalado por D. Kaplan,[50] uno de los autores del estudio crítico que recién comentáramos. Una modificación alternativa ha sido esbozada por J. Kim.[51]

La parte crucial de la revisión de Kim es un requisito que debe agregarse a aquellos especificados en (6.8) a los efectos siguientes: dése a C la forma normal conjuntiva completa en aquellas frases atómicas que esencialmente ocurren en C; entonces, ninguno de los miembros de esa forma normal debe ser lógicamente derivable de E. En nuestra ilustración del primero de los cinco teoremas críticos, se viola este requisito, ya que "Ha" implica lógicamente a cada uno de los miembros de la forma normal con-

[49] R. Eberle, D. Kaplan y R. Montague: "Hempel and Oppenheim on Explanation". *Philosophy of Science,* 28 (1961), págs. 418-28.

[50] D. Kaplan: "Explanation Revisited", *Philosophy of Science,* 28 (1961), págs. 429-36.

[51] J. Kim: "Discussion: On the Logical Conditions of Deductive Explanation". *Philosophy of Science,* 30 (1963), págs. 286-91.

juntiva completa de "$(Fb \vee \sim Ga) \supset Ha$", o sea, "$Fb \vee Ga \vee Ha$", "$\sim Fb \vee Ga \vee Ha$" y "$\sim Fb \vee \sim Ga \vee Ha$". Kim demuestra en general que este requisito adicional bloquea las pruebas ofrecidas por Eberle, Kaplan y Montague para los cinco teoremas que "trivializan" las definiciones (6.8) y (6.9). Sin embargo, sería conveniente cerciorarse con mayor claridad de hasta qué punto el requisito adicional se justifica, no sobre la base *ad hoc* de que bloquea esas pruebas, sino en función de la racionalidad de la explicación científica.

Kaplan enfoca el problema formulando tres requisitos muy plausibles de adecuación para cualquier análisis del tipo explicación deductiva que consideramos aquí. Luego demuestra que el análisis propuesto en la Parte III no satisface conjuntamente aquellos requisitos y que las dificultades exhibidas en los cinco teoremas trivializantes están ligadas a esta deficiencia. Finalmente, revisa las definiciones que se ofrecen en la Parte III a fin de que llenen los requisitos de adecuación y eviten las dificultades que hemos mencionado. Para mayores detalles de esta esclarecedora contribución, el lector deberá consultar el artículo de Kaplan.

BIBLIOGRAFIA

Beard, Charles A. y Sidney Hook: "Problems of Terminology in Historical Writing". Cap. IV de *Theory and Practice in Historical Study: A Report of the Committee on Historiography.* Nueva York, Social Science Research Council, 1946.

Bergmann, Gustav: "Holism, Historicism, and Emergence", *Philosophy of Science,* 11 (1914), 209-221.

Bonfante, G.: "Semantics, Language". En P. L. Harriman (comp.), *The Encyclopedia of Psychology.* Nueva York, 1946.

Broad, C. D.: *The Mind and its Place in Nature.* Nueva York, 1925.

Carnap, Rudolf: *Introduction to Semantics.* Cambridge, Mass., 1942.

Carnap, Rudolf: "On Inductive Logic", *Philosophy of Science,* 12 (1945), 72-97.

Carnap. Rudolf: "The Two Concepts of Probability", *Philosophy and Phenomenological Research,* 5 (1945), 513-532.

Carnap, Rudolf: "Remarks on Induction and Truth", *Philosophy and Phenomenological Research,* 6 (1946), 590-602.

Carnap, Rudolf: "On the Application of Inductive Logic", *Philosophy and Phenomenological Research,* 8 (1947), 133-47.

Chisholm, Roderick M.: "The Contrary-to-Fact Conditional", *Mind,* 55 (1946), 289-307.

Church, Alonzo: "Logic, formal", en Dagobert D. Runes (comp.), *The Dictionary of Philosophy.* Nueva York, 1942.

Ducasse, C. J.: "Explanation, Mechanism, and Teleology", *The Journal of Philosophy,* 22 (1925), 150-55.

Feigl, Herbert: "Operationism and Scientific Method", *Psychological Review,* 52 (1945), 250-59 y 284-88.

Goodman, Nelson: "A Query on Confirmation", *The Journal of Philosophy,* 43 (1946), 383-85.

Goodman, Nelson: "The Problem of Counterfactual Conditionals", *The Journal of Philosophy,* 44 (1947), 113-28.

Goodman, Nelson: "On Infirmities of Confirmation Theory", *Philosophy and Phenomenological Research,* 8 (1947), 149-51.

Grelling, Kurt y Paul Oppenheim: "Der Gestaltbegriff im Lichte der neuen Logik", *Erkenntnis*, 7 (1937-38), 211-25 y 357-59.
Grelling, Kurt y Paul Oppenheim: "Logical Analysis of Gestalt as 'Functional Whole' ". Preprinted for distribution at Fifth Internat. Congress for the Unity of Science, Cambridge, Mass., 1939.
Helmer, Olaf y Paul Openheim: "A Syntactical Definition of Probability and of Degree of Confirmation", *The Journal of Symbolic Logic*, 10 (1945), 25-60.
Hempel, Carl G.: "The Function of General Laws in History". *The Journal of Philosophy*, 39 (1942), 35-48.
Hempel, Carl G.: "Studies in the Logic of Confirmation", *Mind*, 54 (1945); Parte I: págs. 1-26; Parte II: págs. 97-121.
Hempel, Carl G. y Paul Oppenheim: "A Definition of Degree of Confirmation", *Philosophy of Science*, 12 (1945), 98-115.
Henle, Paul: "The Status of Emergence", *The Journal of Philosophy*, 39 (1942), 486-93.
Hospers, John: "On Explanation", *The Journal of Philosophy*, 43 (1946), 337-56.
Hull, Clark L.: "The Problem of Intervening Variables in Moral Behavior Theory", *Psychological Review*, 50 (1943), 273-91.
Hull, Clark L.: *Principles of Behavior.* Nueva York, 1943.
Jevons, W. Stanley: *The Principles of Science.* Londres, 1924 (1ª ed., 1874).
Kaufmann, Felix: *Methodology of the Social Sciences.* Nueva York, 1944.
Knight, Frank H.: "The Limitations of Scientific Method in Economics", en R. Tugwell (comp.), *The Trend of Economics.* Nueva York, 1924.
Koch, Sigmund: "The Logical Character of the Motivation Concept", *Psychological Review*, 48 (1941); Parte I: págs. 15-38; Parte II: págs. 127-154.
Langford, C. H.: Reseña crítica en *The Journal of Symbolic Logic*, 6 (1941), págs. 67-68.
Lewis, C. I.: *An Analysis of Knowledge and Valuation.* La Salle, Ill., 1946.
McKinsey, J. C. C.: Comentario sobre Helmer y Oppenheim (1945). *Mathematical Reviews*, 7 (1946), 45.
Mill, John Stuart: *A System of Logic.* Nueva York, 1858.
Morgan, C. Lloyd: *Emergent Evolution.* Nueva York, 1923.
Morgan, C. Lloyd: *The Emergence of Novelty.* Nueva York, 1933.
Popper, Karl: *Logik der Forschung.* Viena, 1935.
Popper, Karl: *The Open Society and its Enemies.* Londres, 1945.
Reichenbach, Hans: *Elements of Symbolic Logic.* Nueva York, 1947.
Rosenblueth, A., N. Wiener y J. Bigelow: "Behavior, Purpose, and Teleology", *Philosophy of Science*, 10 (1943), 18-24.
Stace, W. T.: "Novelty, Indeterminism and Emergence", *Philosophical Review*, 48 (1939), 296-310.
Tarski, Alfred: "The Semantical Conception of Truth, and the Foundations of Semantics", *Philosophy and Phenomenological Research*, 4 (1944), 341-76.
Tolman, Edward Chase: *Purposive Behavior in Animals and Men.* Nueva York, 1932.
White, Morton G.: "Historical Explanation". *Mind*, 52 (1943), 212-29.
Woodger, J. H.: *Biological Principles.* Nueva York, 1929.
Zilsel, Edgar: *Problems of Empiricism.* Chicago, 1941.
Zilsel, Edgar: "Physics and the Problem of Historico-Sociological Laws", *Philosophy of Science*, 8 (1941), 567-79.

Capítulo XI

La lógica del análisis funcional [1]

1. *Introducción*

La ciencia empírica, en todas sus ramas principales, trata no sólo de *describir* los fenómenos del mundo de nuestra experiencia, sino también de *explicarlos* o *entenderlos*. Si bien esto está ampliamente reconocido, se suele sostener sin embargo que existen diferencias fundamentales entre los *métodos* explicativos adecuados para los diferentes campos de la ciencia empírica. De acuerdo con esta opinión, en las ciencias físicas toda explicación puede obtenerse en último término por referencia a antecedentes causales o correlacionales, mientras que en psicología y en las disciplinas sociales e históricas —según algunos también en biología— el establecimiento de conexiones causales o correlacionales, aun cuando fuera deseable e importante, no es suficiente. Se dice que la comprensión adecuada de los fenómenos estudiados en estos campos requiere otro tipo de explicación.

Uno de los métodos explicativos que se han desarrollado para este propósito es el del análisis funcional, que tan extendido uso ha encontrado en biología, psicología, sociología y antropología. Este procedimiento plantea problemas de considerable interés para la metodología comparativa de la ciencia empírica. El presente ensayo es un intento de esclarecer algunos de estos problemas; su objeto es examinar la estructura lógica del análisis funcional y su significado explicativo y predictivo, por medio de una confrontación explícita con las principales características de los procedimientos explicativos empleados en las ciencias físicas. Comenzaremos pues con un breve examen de estas últimas.

2. *Explicación nomológica: deductiva e inductiva*

En un recipiente lleno de agua hasta el borde a temperatura ambiente flota un trozo de hielo que sobresale por encima de la superficie. A medida que el hielo gradualmente se derrite podría esperarse que el agua desborde

[1] Este artículo se reimprime con algunas modificaciones bajo licencia de Llewellyn Gross (comp.), *Symposium on Sociological Theory*, Nueva York, Harper & Row, 1959.

el recipiente. En la realidad, el nivel del agua permanece invariable. ¿Cómo puede explicarse esto? La clave de la respuesta está dada por el principio de Arquímedes, según el cual un cuerpo sólido que flota en un líquido desplaza un volumen de líquido que tiene igual peso que el cuerpo mismo. De allí que el trozo de hielo tenga el mismo peso que el volumen de agua que su porción sumergida desplaza. Como la fusión no afecta los pesos implicados, el agua en la que el hielo se transforma tiene un peso igual al del hielo y, por lo tanto, el mismo peso que el agua inicialmente desplazada por la porción sumergida del hielo. Puesto que tiene el mismo peso, también tiene el mismo volumen que el agua desplaza; por lo tanto la fusión del hielo produce un volumen de agua que basta para llenar con exactitud el espacio inicialmente ocupado por el trozo sumergido del hielo. En consecuencia, el nivel del agua permanece invariable.

Este relato (que deliberadamente descarta ciertos efectos de escasa magnitud) ilustra un razonamiento que intenta explicar un hecho dado. Como cualquier razonamiento explicativo consta de dos partes que serán denominadas *explanans* y *explanandum*.[2] Este último es el enunciado o conjunto de enunciados que describe al fenómeno que ha de explicarse; el primero es el enunciado o conjunto de enunciados que se aducen para proporcionar una explicación. En nuestro ejemplo, el explanandum afirma que al final del proceso el recipiente sólo contiene agua, cuya superficie tiene el mismo nivel que al comienzo. Para explicar este hecho el explanans aduce, en primer término, ciertas leyes de la física (entre ellas el principio de Arquímedes) que señalan que a temperaturas superiores a cero grado y a presión atmosférica, un trozo de hielo se convierte en una masa de agua de igual peso; y la ley de que a temperatura y presión constantes, cantidades de agua iguales en peso también son iguales en volumen.

Además de estas leyes el explanans contiene un segundo grupo de enunciados que describen ciertas circunstancias particulares previas al resultado que quiere explicarse en el experimento, tales como el hecho de que al comienzo hay un trozo de hielo que flota en un recipiente lleno de agua; que el agua se halla a temperatura ambiente y que el recipiente está rodeado de aire a igual temperatura, inalterable hasta el fin del experimento.

El alcance explicativo del razonamiento total reside en demostrar que el resultado descripto en el explanandum era de esperarse en vista de las

[2] Preferimos estos términos en lugar de las conocidas palabras "explicans" y "explicandum" que reservaremos para su utilización en el contexto de la explicación filosófica, tal como lo propone R. Carnap en el sentido técnico. Véase por ejemplo su *Logical Foundations of Probability* (Chicago, University of Chicago Press, 1950), secciones 1-3. Por esta razón se introdujeron los términos "explanans" y "explanandum" en un artículo anterior: Carl Hempel y P. Oppenheim: "Studies in the Logic of Explanation". *Philosophy of Science*, 15, 1948, págs. 135-75 (incluido en el presente volumen). Si bien ese artículo no se ocupa explícitamente de la explicación inductiva, sus cuatro primeras secciones contienen otras consideraciones sobre la explicación deductiva, relevantes al presente estudio. Para un examen crítico de algunos detalles tratados en el artículo anterior, en particular la relación entre explicación y predicción, véase el trabajo de I. Scheffler: "Explanation, Prediction, and Abstraction". *The British Journal of the Philosophy of Science*, 7, 1957, págs. 293-309, que también contiene algunos comentarios interesantes relacionados con el análisis funcional.

circunstancias antecedentes y de las leyes generales numeradas en el explanans. Con mayor precisión, la explicación puede construirse como un razonamiento en que el explanandum se deduce del explanans. Nuestro ejemplo ilustra, pues, lo que llamaremos explicación por subsunción deductiva bajo leyes generales o, brevemente, *explicación nomológica deductiva.* La forma general de tal explicación está dada por el siguiente esquema:

$$(2.1) \qquad \left.\begin{array}{c} L_1, L_2, \ldots, L_m \\ C_1, C_2, \ldots, C_n \end{array}\right\} \text{Explanans}$$

$$\overline{\qquad E \qquad} \quad \text{Explanandum}$$

Aquí, L_1, L_2, ..., L_m son leyes generales, y C_1, C_2, ..., C_n enunciados de hechos particulares; la línea horizontal que separa la conclusión E de las premisas indica que aquélla se desprende lógicamente de las últimas.

En nuestro ejemplo, el fenómeno a explicarse es un hecho concreto que ocurre en un cierto lugar y tiempo. Pero el método de la subsunción deductiva bajo leyes generales se presta también a la explicación de lo que puede denominarse "hechos generales" o regularidades, tales como las expresadas en leyes de la naturaleza. Por ejemplo, la pregunta de por qué la ley de Galileo es válida para cuerpos físicos que caen libremente cerca de la superficie terrestre, puede responderse demostrando que esta ley se refiere a un caso especial de movimiento acelerado por la atracción de la gravedad, y que puede deducirse de las leyes generales para se movimiento (es decir, las leyes del movimiento y de la gravedad de Newton) aplicándolas al caso especial en que estén implicados dos cuerpos, uno de los cuales es la Tierra y el otro el objeto que cae, y donde la distancia entre sus centros de gravedad es igual a la longitud del radio terrestre. Así puede lograrse una explicación de las regularidades expresas por la ley de Galileo, deduciendo la última de las leyes newtonianas y de enunciados que especifican la masa y el radio de la Tierra; estas dos últimas proporcionan el valor de la constante de aceleración de caída libre cercana a la Tierra.

Puede ser útil mencionar otro ejemplo del papel de la explicación nomológica deductiva para ilustrar hechos particulares, como también de uniformidades generales o leyes. La formación de un arco iris en un momento dado puede explicarse deductivamente por referencia a: 1) ciertas condiciones determinantes, tales como la presencia de gotas de lluvia en el aire, la luz solar que incide sobre estas gotas, el observador que está de espaldas al sol, etc., y 2) ciertas leyes generales, especialmente aquellas de la reflexión, refracción y dispersión ópticas. El hecho de que estas leyes sean válidas puede explicarse a su vez por deducción a partir de principios más inclusivos, por ejemplo de la teoría electromagnética de la luz.

Así, el método de la explicación nomológica deductiva describe un hecho particular subsumiéndolo bajo leyes generales, tal como queda representado en el esquema (2.1). De un modo similar puede servir para explicar el hecho de que una cierta ley sea válida demostrando que es sub-

sumible de la misma manera bajo leyes o principios teóricos más globales. En realidad, uno de los principales objetivos de una teoría (como la teoría electromagnética de la luz, por ejemplo) es precisamente proveer un conjunto de principios —a menudo expresados en términos de entidades "hipotéticas" no observables (de manera como los vectores de un campo eléctrico o magnético)— que expliquen deductivamente un grupo de "generalizaciones empíricas" antes establecidas (como las leyes de la propagación rectilínea, la reflexión y la refracción de la luz). Con frecuencia, una explicación teórica mostrará que las generalizaciones empíricas sólo tienen validez aproximada. Por ejemplo, la aplicación de la teoría de Newton a la caída libre cerca de la Tierra conduce a una ley similar a la de Galileo, excepto en que la aceleración de la caída no se considera estrictamente constante, sino que varía apenas con la ubicación geográfica, la altura sobre el nivel del mar y ciertos otros factores.

Las leyes generales o principios teóricos que sirven para explicar las generalizaciones empíricas pueden a su vez ser subsumibles deductivamente bajo principios aun más globales: por ejemplo, la teoría de la gravedad de Newton puede ser subsumida, como aproximación, en la teoría general de la relatividad. Es obvio que esta jerarquía explicativa tiene que terminar en algún punto. O sea que en cualquier momento del desarrollo de la ciencia empírica habrá ciertos hechos que no son explicables. Esto incluye las leyes generales y los principios teóricos más globales conocidos hasta entonces y, por supuesto, muchas generalizaciones empíricas y hechos particulares para los cuales no se dispone en ese momento de principios explicativos. Pero esto no quiere decir que ciertos hechos sean intrínsecamente inexplicables y que lo sean para siempre. Cualquier hecho particular hasta ahora inexplicable y cualquier principio general, por muy inclusivo que sea, puede encontrar su explicación posterior por subsunción bajo principios aun más inclusivos.

La explicación causal es un tipo especial de explicación nomológica deductiva. Puede decirse que determinado hecho o conjunto de hechos ha causado un "efecto" específico, sólo si hay leyes generales que conecten el primero con el último, de modo que dada una descripción de los hechos antecedentes, la ocurrencia del efecto puede deducirse con la ayuda de las leyes. Por ejemplo, la explicación de que la dilatación de una barra de hierro ha sido causada por un aumento de su temperatura nos lleva a un razonamiento de la forma (2.1), cuyo explanans incluye (a) enunciados que especifican la longitud inicial de la barra, indicando que la barra es de hierro y que se ha aumentado su temperatura; (b) una ley que establece que la longitud de una barra de hierro aumenta con la temperatura.[3]

Sin embargo, no toda explicación nomológica deductiva es una ex-

[3] Una explicación por medio de leyes que son causales, en el sentido técnico de la física teórica tiene también la forma (2.1) de una explicación nomológica deductiva. En este caso, las leyes deben reunir ciertas condiciones en cuanto a la forma matemática y C_1, C_2, ..., C_n expresar las condiciones límite. Para una descripción más completa de los conceptos de leyes causales y causalidad como se los entiende en la física teórica véase, por ejemplo, H. Margenau: *The Nature of Physical Reality.* Nueva York, McGraw-Hill Book Company, Inc., 1950, capítulo 19; o bien Ph. Frank: *Philosophy of Science,* Englewood Cliffs, N.J., Prentice-Hall, Inc., capítulos 11, 12.

plicación causal. No podemos decir con propiedad, por ejemplo, que las regularidades expresadas por las leyes del movimiento y de la gravedad de Newton causen la caída libre de cuerpo cercanos a la superficie de la Tierra para cumplir con las leyes de Galileo.

Debemos considerar, pues, otro tipo de explicación que nuevamente explique un fenómeno dado, refiriéndolo a leyes generales de un modo que no se adecua al esquema deductivo (2.1). Cuando un niño, Enrique, se enferma de paperas, podemos explicarlo diciendo que contrajo la enfermedad al contagiarse de un amigo con el cual jugó un par de horas un día antes que este último cayera enfermo. Los antecedentes específicos aquí invocados son: que Enrique estuvo expuesto y, además, según suponemos, que no se hallaba ya afectado por la enfermedad.

Pero para relacionar esto con el hecho que debe explicarse, no podemo aducir una ley general en el sentido de que bajo las condiciones mencionadas, la persona invariablemente contrae paperas. Lo único que puede afirmarse es que la enfermedad tiene una alta probabilidad estadística de ser transmitida. De un modo similar, cuando un rasgo neurótico en un adulto se puede explicar desde un punto de vista psicoanalítico por referencia a experiencias infantiles críticas, este razonamiento sostiene implícita o explícitamente que el caso considerado no es más que un ejemplo de ciertas leyes generales que gobiernan el desarrollo de las neurosis. Cualesquiera que sean las leyes específicas de este tipo que puedan aducirse hoy en día expresarán, en el mejor de los casos, más bien tendencias probabilísticas que regularidades determinísticas. Se las puede construir como *leyes de forma estadística* o, dicho con más brevedad, como *leyes estadísticas* con el fin de que, dadas las experiencias infantiles en cuestión —además de, presumiblemente, ciertas condiciones ambientales posteriores de la vida—, existe tal y tal probabilidad estadística de que se desarrolle un tipo específico de neurosis. Estas leyes estadísticas difieren en su forma de las leyes universales mencionadas en nuestros anteriores ejemplos sobre razonamientos explicativos. En el caso más simple una *ley de forma estrictamente universal,* o mejor, una *ley universal,* es un enunciado por el cual en *todos* los casos que cumplan ciertas condiciones antecedentes *A* (por ejemplo, el calentamiento de un gas a presión constante) se producirá un hecho específico B (por ejemplo, un aumento del volumen del gas), mientras que una ley de forma estadística afirma que la probabilidad para que las condiciones *A* estén acompañadas de un hecho *B* tiene un valor específico *p.*

Los razonamientos explicativos que dan razón de un fenómeno, como recién se ilustró, por referencia a leyes estadísticas, no son del tipo estrictamente deductivo (2.1). Por ejemplo: el explanans constituido por la información acerca de que Enrique estuvo expuesto a las paperas y de una ley estadística sobre la transmisión de la enfermedad no implica, desde un punto de vista lógico, la conclusión de que el niño se contagie de paperas. Por ello no se pretende que la conclusión sea necesaria, pero podríamos decir que es más o menos probable, dependiendo de la probabilidad especificada por las leyes estadísticas. O sea, que un razonamiento de este tipo explica un fenómeno, demostrando que su ocurrencia es altamente probable dados determinados hechos y leyes estadísticas espe-

cificadas en el explanans. Una explicación de este tipo se llamará *explicación por subsunción bajo leyes estadísticas* o bien, en forma más breve, *explicación inductiva.*

Un análisis más detallado muestra que la explicación inductiva difiere de su contraparte deductiva en ciertos aspectos importantes; [4] pero para el propósito de la discusión que sigue será suficiente la descripción esquemática de la explicación por leyes estadísticas.

Acerca de ambos tipos de explicación que hemos distinguido diremos que son formas de *explicación nomológica,* ya que cualquiera de ellas explica un fenómeno dado por medio de su "subsunción bajo leyes", es decir, demostrando que su ocurrencia podría inferirse —deductivamente o con una alta probabilidad— aplicando ciertas leyes de forma universal o estadística a circunstancias antecedentes especificadas. Por eso, una explicación nomológica muestra que en realidad se podría haber *predicho* el fenómeno en cuestión sea deductivamente o con alta probabilidad si, en un momento anterior, hubiéramos tenido conocimiento de los hechos establecidos en el explanans.

Sin embargo, el poder predictivo de la explicación nomológica tiene mucho mayor alcance: precisamente porque su explanans contiene leyes generales permite las predicciones correspondientes a sucesos distintos de los referidos en el explanandum. En realidad estas predicciones proporcionan los medios para probar la verdad empírica del explanans. Por ejemplo: las leyes invocadas en una explicación deductiva de la forma (2.1) implican que el tipo de suceso que se describe en E se repetirá cada vez que se presenten circunstancias del tipo descripto por $C_1, C_2, \ldots, C_n$. Es decir, que cuando se repita el experimento del hielo que flota en el agua el resultado será el mismo. Más aún, las leyes harán predicciones sobre lo que ocurrirá en ciertas condiciones especificables que difieran de aquellas mencionadas en $C_1, C_2, \ldots, C_n$. Así, las leyes invocadas en nuestro ejemplo también podrán predecir que si un trozo de hielo flotara en un recipiente lleno hasta el borde con salmuera concentrada (que tiene mayor peso específico que el agua) parte del líquido desbordaría a medida que el hielo se fuera derritiendo. Aquí también las leyes del movimiento y de la gravedad de Newton, que pueden usarse para explicar distintos aspectos del movimiento planetario, tienen consecuencias predictivas para una serie de fenómenos totalmente diferentes como la caída libre cerca de la Tierra, el movimiento de un péndulo, las mareas y muchos otros.

Este tipo de predicción de otros fenómenos posibilitados por la explicación nomológica no se limita a sucesos futuros; también puede refe-

[4] Para mayores detalles véase la sección 3 del capítulo sobre "Aspectos de la explicación científica" en el presente volumen. Algunos comentarios lúcidos sobre la explicación por medio de leyes estadísticas pueden consultarse en S. E. Gluck: "Do Statistical Laws have Explanatory Efficacy?" *Philosophy of Science,* 22 (1955), págs. 34-38. Para un análisis más completo de la lógica de la inferencia estadística, véase R. B. Braithwaite: *Scientific Explanation.* Cambridge, Cambridge University Press, 1953, capítulos V, VI, VII. Para un estudio sobre la lógica de la inferencia deductiva en general, la obra de Carnap, *Logical Foundations of Probability, op. cit.,* es de suma importancia.

rirse a los del pasado. Dada, por ejemplo, cierta información sobre la actual posición y velocidades de los cuerpos celestes, los principios de la mecánica newtoniana y de la óptica no sólo arrojarán predicciones sobre los eclipses solares y lunares futuros, sino también "posdicciones" o "retrodicciones" de los pasados. De un modo análogo, las leyes estadísticas sobre el decaimiento radioactivo que pueden funcionar en diversos tipos de predicciones también se prestan al uso retrodictivo, por ejemplo para fechar, por medio del método de radiocarbonos, el arco o el mango del hacha que se encuentran en un yacimiento arqueológico.

Una explicación propuesta sólo es aceptable científicamente si su explanans se presta a la prueba empírica, es decir, si es posible inferir de aquél ciertos enunciados cuya verdad pueda verificarse por medio de adecuados procedimientos experimentales o de observación. Las implicaciones predictivas y posdictivas de las leyes invocadas en una explicación nomológica ofrecen claramente una oportunidad para pruebas empíricas. Cuanto más variado y extenso sea el conjunto de implicaciones que han sido confirmadas por la investigación empírica, mejor establecidos estarán los principios explicativos en cuestión.

3. *Pauta básica del análisis funcional*

Desde un punto de vista histórico, el análisis funcional es una modificación de la explicación teleológica, es decir, de la explicación no tanto por referencia a causas que "producirán" el hecho en cuestión, sino por referencia a fines que determinan su curso. Intuitivamente, parece muy plausible que para una comprensión adecuada de la conducta finalista y de otras orientadas a una meta, se requiera un enfoque teleológico; en ese contexto la explicación teleológica siempre ha tenido sus defensores. La dificultad con la idea reside en que en sus formas más tradicionales no satisfacen los requisitos científicos mínimos de la prueba empírica. La idea neovitalista de entelequia o de fuerza vital es un ejemplo. Se propone que ésta nos proporcionará una explicación para varios fenómenos biológicos característicos, tales como la regeneración y la regulación, los que de acuerdo con el neovitalismo no pueden explicarse tan sólo por las leyes físicas y químicas. Las entelequias se conciben como agentes no físicos orientados a una meta y que afectan el curso de los hechos fisiológicos, de modo que restituyen el organismo a un estado más o menos normal después de producirse la perturbación. Sin embargo, esta concepción se enuncia en términos esencialmente metafóricos: no se suministra un conjunto testable de enunciados (a) que especifiquen los tipos de circunstancias en que sobrevendrá una entelequia como agente que dirige el curso de los acontecimientos gobernados por leyes físicas y químicas, y (b) que indique con precisión qué efectos observables tendrá la acción de una entelequia en un caso así. Como el neovitalismo no puede establecer leyes generales acerca de cómo y cuándo actúa una entelequia, no puede explicar ningún fenómeno biológico. No nos da fundamentos para que un fenómeno sea esperable, ningún motivo para decir:

"Ahora vemos que el fenómeno tenía que ocurrir". No permite ni las predicciones ni las retrodicciones: la atribución de un fenómeno biológico a la aparición de una entelequia no tiene ninguna implicación testable. Este defecto teórico se pone de relieve al contrastar la idea de entelequia con la de un campo magnético generado por una corriente eléctrica, que puede invocarse para explicar la desviación de una aguja magnética. Tampoco se puede observar de manera directa un campo magnético como una entelequia. Sin embargo, el concepto es gobernado por leyes estrictamente especificables que se refieren a la fuerza y dirección, en cualquier punto, del campo magnético producido por una corriente que circula a través de un conductor dado, además de otras leyes que determinan el efecto de este campo sobre la aguja magnética en el campo magnético sobre la Tierra. Precisamente estas leyes son las que por su importancia predictiva y retrodictiva confieren poder explicativo al concepto de campo magnético. Las descripciones teleológicas que se refieren a entelequias son, pues, seudoexplicaciones. Como se verá, el análisis funcional, si bien a menudo se formula en términos teleológicos, no necesita apelar a esas entidades problemáticas y tiene un núcleo decididamente empírico.

El tipo de fenómeno que el análisis funcional[5] pretende explicar es alguna actividad recurrente o pauta de conducta en un individuo o grupo, tal como un mecanismo fisiológico, un rasgo neurótico, una pauta cultural o institución social. El objetivo principal del análisis es mostrar la contribución que realiza la pauta de conducta a la preservación o desarrollo del individuo o del grupo en que se presenta. Por lo tanto, el análisis funcional trata de comprender una pauta de conducta o una institución sociocultural, determinando el rol que desempeñan para mantener en buen funcionamiento un sistema dado o para que siga siendo un proyecto viable.

A modo de ilustración simple y esquematizada considérese, en primer lugar, el enunciado:

(3.1) El latido del corazón en los vertebrados tiene como función hacer circular la sangre por el organismo.

Antes de examinar las posibilidades de uso como explicación deberíamos preguntarnos: ¿Qué *significa* este enunciado? ¿Qué se afirma con esta atribución de función? Podría alegarse que toda la información transmitida por un enunciado como (3.1) puede también expresarse sustiutyendo la palabra "función" por "efecto". Pero esta acepción nos obliga a aceptar también el enunciado:

(3.2) El latido del corazón tiene la función de producir sonidos cardíacos, porque tiene ese efecto.

[5] Para caracterizar el análisis funcional presentado en esta sección he recibido mucho estímulo e información del interesante ensayo "Manifest and Latent Functions" en el libro de R. K. Merton: *Social Theory and Social Structure*. Nueva York, The Free Press, edición corregida y aumentada, 1957, 19-84. Cada uno de los párrafos a los que se refiere el presente trabajo pueden encontrarse también en la primera edición (1949) cuyo número de página es casi igual.

Sin embargo, un propulsor del análisis funcional se rehusaría a afirmar (3.2) debido a que los sonidos cardíacos son un efecto del latido que no tiene ninguna importancia para el funcionamiento del organismo. Por otra parte, la circulación de la sangre realiza el transporte de material nutritivo y la eliminación de residuos de distintas partes del cuerpo: proceso que es indispensable para que el organismo continúe funcionando e incluso para que viva. Entendido de esta maneera, la importancia del enunciado funcional (3.1) podría resumirse así:

(3.3) El latido del corazón tiene el efecto de hacer circular la sangre y esto asegura la satisfacción de ciertas condiciones (provisión de material nutritivo y eliminación de residuos) que son necesarias para el adecuado funcionamiento del organismo.

Ahora debemos destacar que el corazón cumplirá las funciones que le hemos atribuido sólo si el organismo y el ambiente logran satisfacer ciertas condiciones. Por ejemplo, si la aorta presenta una ruptura la circulación fallará; la sangre sólo puede transportar oxígeno si el medio le proporciona una adecuada provisión y si los pulmones están en buenas condiciones; eliminará ciertos residuos sólo si los riñones están razonablemente sanos, etc. La mayoría de las condiciones aquí especificadas no se mencionan en general, en parte porque sin duda se supone que se cumplen regularmente en situaciones de normalidad del organismo. Esta omisión refleja la falta de conocimiento adecuado, ya que una especificación explícita de condiciones relevantes exigiría una teoría en la que (a) los posibles estados de los organismos y sus medios pudieran caracterizarse por valores de ciertas "variables de estado" fisicoquímicas o tal vez biológicas, y (b) en la cual los principios teóricos fundamentales permitieran la determinación de esa gama de condiciones internas y externas en las que las pulsaciones del corazón cumplieran su función, como ya lo hemos referido.[6] Por el momento no disponemos, naturalmente, de una teoría general de este tipo y ni siquiera de una teoría que se ocupe de ese modo de alguna clase de organismos.

Además, una reformulación completa de (3.1) a la manera de (3.3) exige criterios acerca de qué es "desempeño adecuado", "funcionamiento normal", etc., del organismo en cuestión. Porque la función de un rasgo dado se pondera aquí en función de su relevancia causal respecto de satisfacer ciertas condiciones necesarias para el adecuado funcionamiento o la supervivencia del organismo. También aquí los criterios requeridos a menudo quedan sin especificar, aspecto éste del análisis funcional cuyas graves implicaciones trataremos más adelante (en la sección 5).

Las consideraciones aquí esbozadas sugieren la siguiente caracterización esquemática de un análisis funcional:

(3.4) *Pauta básica de un análisis funcional*: El objeto del análisis

[6] Para una exposición completa y mejor desarrollo de este punto véase "A Formalization of Functionalism" en E. Nagel: *Logic Without Metaphysics*. Nueva York, The Free Press, 1957, págs. 247-83. La parte 1 de esa investigación ofrece un análisis detallado del trabajo de Merton mencionado en la nota 5.

es un "ítem" *i* que constituye un rasgo o disposición relativamente persistente (por ejemplo, los latidos del corazón) que ocurre en un sistema *s* (verbigracia, el cuerpo de un vertebrado vivo). El análisis se propone demostrar que *s* se encuentra en un estado o condición interna c_i y en un medio que presenta ciertas condiciones externas c_e tales, que en condiciones c_i y c_e (juntas las denominaremos *c*) el rasgo tenga efectos que satisfagan alguna "necesidad" o "requerimiento funcional" de *s*, es decir, una condición *n* que es necesaria para que el sistema siga funcionando de manera adecuada, eficiente o regular. Pasemos a considerar con brevedad algunos ejemplos de este tipo de análisis en psicología y en los estudios sociológicos y antropológicos. En psicología es primordialmente el psicoanálisis el que presenta una fuerte orientación funcional. Un ejemplo claro es la caracterización funcional que hace Freud del rol de la formación de síntomas. En *El problema de la ansiedad*, Freud expresa su acuerdo con una concepción por la cual "toda formación de síntomas se produce exclusivamente para evitar la ansiedad; los síntomas ligan la energía psíquica que de otro modo se descargaría como ansiedad".[7] Para apoyar esta noción Freud señala que si a un agorafóbico que normalmente es acompañado para salir se lo deja solo en la calle, sufrirá un ataque de ansiedad, igual que el neurótico compulsivo a quien se le impide lavarse las manos después de haber tocado algo. "Resulta evidente, pues, que la estipulación de estar acompañado y la compulsión a lavarse tiene su propósito y también su resultado: evitar el ataque de ansiedad."[8] En este relato, formulado en términos fuertemente teleológicos, el sistema *s* es el individuo considerado; *i* es la pauta de conducta agorafóbica o compulsiva; *n* la inhibición de la angustia que es necesaria para evitar una crisis psicológica grave que haría imposible que el individuo funcione adecuadamente.

En antropología y sociología el objeto del análisis funcional, según las palabras de Merton, es "un ítem *estandardizado* (es decir, normalizado y repetitivo como los roles sociales, las pautas institucionales, los procesos sociales, las pautas culturales, las emociones culturales conformadas, las normas sociales, la organización de grupo, la estructura social, los mecanismos de control social, etc.)".[9] Tanto aquí como en psicología y biología la función (es decir, el efecto estabilizador o de ajuste) del ítem estudiado puede ser conscientemente buscada o no por los agentes (a veces ni siquiera será reconocida de moodo consciente). En este caso, Merton habla de funciones *latentes*, distinguiéndolas de las funciones *manifiestas*, es decir, aquellos efectos objetivos de estabilización que se proponen los participantes de un sistema.[10] Así, por ejemplo, las ceremonias de los

[7] S. Freud: *The Problem of Anxiety* (Traducción de H. A. Bunker.) Nueva York, Psychoanalytic Quarterly Press y W. W. Norton & Company, Inc., 1936, pág. 111.

[8] *Ibíd.*, pág. 112.

[9] Merton, *obra citada*, pág. 50 (bastardilla del autor).

[10] *Ibíd.*, pág. 51. Merton define a las funciones manifiestas como aquellas que son a la vez intencionales y reconocidas; a las funciones latentes como aquellas que no son ni intencionales ni reconocidas. Sin embargo, esta caracterización da lugar a funciones que no son ni manifiestas ni latentes, por ejemplo aquellas que son reconocidas pero no intencionales. Por lo tanto, parecería estar más en el espíritu de

Hopi para provocar lluvias no logran su objetivo meteorológico manifiesto pero "pueden cumplir la función latente de reforzar la identidad del grupo proporcionando una ocasión periódica para que los miembros dispersos se reúnan y participen de una actividad común".[11]

El análisis funcional de Radcliffe-Brown sobre los ritos totémicos de ciertas tribus australianas ilustra el mismo punto:

> Para descubrir la función social de los ritos totémicos debemos considerar todo el aparato de ideas cosmológicas del cual cada rito es una expresión parcial. Creo que es posible demostrar que la estructura social de una tribu australiana está relacionada, de una manera muy especial, con estas ideas cosmológicas, dado que la conservación de su continuidad depende de mantenerlas siempre vivas por medio de la expresión común en mitos y ritos.
>
> Por lo tanto, cualquier estudio satisfactorio acerca de los ritos totémicos en Australia no debe basarse simplemente en la consideración de su propósito manifiesto..., sino en el descubrimiento de su significado y de su función social.[12]

Malinowski atribuye funciones latentes importantes a la religión y a la magia. Para él la fe religiosa establece y destaca actitudes mentales como el respeto por la tradición, la armonía con el medio ambiente, la seguridad y el coraje en situaciones críticas y frente a la perspectiva de muerte —actitudes que, incorporadas y fomentadas por el culto y el ceremonial tienen "un inmenso valor biológico". Señala que la magia, por proporcionar al hombre ciertos rituales, técnicas y creencias fijas, le permite "mantener su equilibrio e integridad mental en arranques de ira, en las garras del odio, del amor no correspondido, en la desesperación y la ansiedad. La función de la magia es ritualizar el optimismo del hombre, reforzar su fe en la victoria de la esperanza sobre el miedo".[13]

En seguida tendremos oportunidad de agregar a los ejemplos precedentes, tomados del psicoanálisis y la antropología, algunas instancias de análisis funcional en sociología. Sin embargo, para ilustrar el carácter general del procedimiento bastarán los casos mencionados hasta ahora: todos ellos exhiben el esquema básico esbozado en (3.4). Pasaremos ahora

---

Merton basar la distinción según que el efecto estabilizador de un ítem dado, sea deliberadamente buscado o no.

[11] *Ibíd.*, págs. 64-65.

[12] A. R. Radcliffe-Brown: *Structure and Function in Primitive Society.* Londres, Cohen and West Ltd., 1952, pág. 145.

[13] B. Malinowski: *Magic, Science and Religion, and Other Essays.* Garden City, Nueva York, Doubleday Anchor Books, 1954, pág. 90. Una comparación esclarecedora de las opiniones de Malinowski sobre las funciones de la magia y de la religión con las de Radcliffe-Brown, es la de G. C. Homans: *The Human Group.* Nueva York, Hartcourt, Brace & World, Inc., 1950, págs. 321 y sigs. (Léase también los comentarios generales de Homans sobre "la teoría funcional", *Ibíd.* págs. 268-72). Este tema y otros aspectos del análisis funcional en antropología se examinan críticamente en el siguiente artículo que confronta algunas aplicaciones específicas del método con las declaraciones programáticas de sus propulsores: Leon J. Goldstein: "The Logic of Explanation in Malinowskian Anthropology", *Philosophy of Science,* 24 (1957), págs. 156-66.

del examen de la forma del análisis funcional a una apreciación de su alcance como modo de explicación.

4. *El alcance explicativo del análisis funcional*

En general se considera que el análisis funcional logra una *explicación* de los "ítems" cuyas funciones estudia. Malinowski, por ejemplo, sostiene que el análisis funcional de la cultura "se propone explicar los hechos antropológicos en todos sus niveles de desarrollo a través de su función..." [14] y agrega en el mismo contexto: "Explicar cualquier ítem de la cultura material o moral significa señalar su lugar funcional dentro de una institución..."[15] En otro párrafo Malinowski habla de la "explicación funcional del arte, la recreación y las ceremonias públicas".[16]

También Radcliffe-Brown considera que el análisis funcional es un método explicativo, si bien no el único adecuado para las ciencias sociales: "De un modo similar una 'explicación' de un sistema social será su historia hasta donde la conozcamos: un relato detallado de cómo llegó a ser lo que es y dónde está. Otra 'explicación' del mismo sistema puede obtenerse demostrando (como los funcionalistas han tratado de hacer) que se trata de una simplificación especial de las leyes de la fisiología social o del funcionamiento social. Los dos tipos de explicación no entran en conflicto sino que se complementan".[17]

Además de ilustrar la atribución del significado explicativo del análisis funcional, este párrafo es interesante porque destaca que el análisis funcional debe basarse en leyes generales. Lo mismo se demuestra en nuestra caracterización esquemática (3.4): los enunciados de que i en un ámbito específico c tiene efectos que satisfacen a n, siendo n una condición necesaria para el buen funcionamiento del sistema, ambos incluyen leyes generales. Para un enunciado de relación causal esto es muy conocido. La afirmación de que una condición n constituye un prerrequisito funcional para un estado de cierto tipo específico (como el buen funcionamiento) equivale al enunciado de una ley, en el sentido de que cada vez que no se cumple la condición n, no se produce el estado en cuestión. Por lo tanto, la explicación por medio del análisis funcional requiere hacer referencia a leyes.[18]

[14] B. Malinowski: "Anthropology", *Encyclopaedia Britannica,* Primer Volumen Suplementario, Londres y Nueva York; The Encyclopaedia Britannica, Inc., 1926, pág. 132.

[15] *Ibíd.,* pág. 139.

[16] B. Malinowski: *A Scientific Theory of Culture, and other Essays.* Chapel Hill, University of North Carolina Press, 1944, pág. 174.

[17] Radcliffe-Brown, *obra citada,* pág. 186. Para un análisis de la idea de explicación historicogenética a la que se hace referencia en este pasaje, véase la sección 7 del capítulo XII.

[18] En un lugar de su obra, Malinowski respalda una declaración que podría juzgarse como disidente con esta conclusión: "La descripción no puede separarse de la explicación ya que en las palabras de un gran físico 'la explicación no es otra cosa que descripción condensada'" (Malinowski: "Anthropology", *obra citada,* pág. 132). Parecería referirse a las opiniones de Ernst Mach o Pierre Duhem quienes

¿Qué alcance explicativo podemos exigir, pues, al análisis funcional? Supongamos que tenemos interés de explicar la ocurrencia de un rasgo *i* en un sistema *s* (en un momento *t*) y que se brinda el siguiente análisis funcional:

(4.1)

a) En *t*, *s* funciona adecuadamente en un ámbito de tipo *c* (caracterizado por condiciones internas y externas específicas)
b) *s* funciona adecuadamente en un ámbito de tipo *c* sólo si se satisface cierta condición necesaria *n*
c) Si el rasgo *i* estuviera presente en *s*, entonces como efecto, se cumpliría la condición *n*
d) (Por lo tanto) en *t*, el rasgo *i* está presente en *s*

Por el momento dejemos de lado la cuestión acerca de qué significan exactamente enunciados del tipo (a) y (b) y en especial la oración "*s* funciona adecuadamente"; esto se examinará en la sección 5. Ahora nos ocuparemos tan sólo de la *lógica* del razonamiento, es decir, que nos preguntaremos si (d) se desprende formalmente de (a), (b), (c), del mismo modo que en la explicación nomológica deductiva el explanandum se desprende del explanans. Obviamente la respuesta es negativa dado que el razonamiento (4.1) involucra la falacia de afirmar el consecuente en relación con la premisa c). De manera más explícita, sería válido inferir es que la condición *n* se debe cumplir de un modo u otro en el tiempo dría lograr satisfacer la condición *n*. De hecho, lo único que se puede inferir es que la condición *n* se debe cumplir de un modo u otro en el tiempo *t*, porque si no, a causa de b), el sistema *s* no podría funcionar adecuadamente en su ámbito, que es una contradicción de lo que afirma (a). También podría ser que si ocurriera alguno de una serie de ítems alternativos, bastaría tanto como *i*, para satisfacer el requisito *n*, en cuyo caso la explicación dada por las premisas de (4.1) simplemente omite decir por qué el rasgo *i* está presente en *s* en el momento *t* y no alguna de sus alternativas.

Como se ha señalado, esta objeción no se aplicaría si la premisa (c)

---

defendieron una posición similar sobre este punto. Mach concebía como objetivo básico de la ciencia la descripción breve y sucinta de los fenómenos recurrentes, y consideraba las leyes como una manera muy eficiente de comprimir la descripción de una infinidad de sucesos potenciales en una fórmula simple y compacta. Entendida de esta manera, la oración citada por Malinowski, con la cual concuerda, es por supuesto totalmente compatible con nuestra afirmación sobre la relevancia de las leyes para la explicación funcional.

Además, una ley sólo puede denominarse descripción en un sentido muy amplio. Incluso una generalización tan simple como "Todos los vertebrados tienen corazón" no describe a un individuo en particular, tal como Rin-Tin-Tin, que es un vertebrado y que posee un corazón. Lo que en todo caso enuncia sobre Rin-Tin-Tin y cualquier otro objeto, vertebrado o no, es que *si* es un vertebrado *entonces* poseerá corazón. Por lo tanto, la generalización tiene al alcance de un conjunto indefinido de enunciados condicionales sobre objetos particulares. Además, podría decirse que una ley implica razonamientos sobre "sucesos potenciales" que en la realidad no se producen. La ley de los gases, por ejemplo, implica que si se calentara un gas a presión constante en el momento *t*, su volumen aumentaría. Pero si realmente el gas no se calienta en *t* este enunciado no puede invocarse como la descripción de algún hecho particular.

pudiera reemplazarse por el enunciado de que el requisito *n sólo* puede ser llenado por la presencia del rasgo *i*. En efecto, algunas instancias del análisis funcional parecerían indicar la pretención de que el ítem específico que se está analizando es, en este sentido, funcionalmente indispensable para el cumplimiento de *n*. Malinowski, por ejemplo, lo pretende para la magia cuando afirma que ésta "desempeña una función indispensable dentro de la cultura: satisface una necesidad definida que no puede ser satisfecha por ninguno de los otros factores de la civilización primitiva". Lo mismo cuando dice de la magia que "sin su poder y conducción, el hombre primitivo no podría haber dominado sus dificultades prácticas como lo hizo, ni podría haber progresado hacia las etapas superiores de la cultura. De ahí la aparición universal de la magia en las sociedades primitivas y su enorme influjo. De ahí que encontremos la magia como complemento invariable de todas las actividades importantes".[19]

Sin embargo, suponer que un ítem dado es funcionalmente indispensable se presta a un fuerte cuestionamiento sobre bases empíricas: en todos los casos concretos de aplicación no parecen existir alternativas. Por ejemplo, la represión de la ansiedad en un sujeto dado puede realizarse con un síntoma alternativo, como lo confirma la experiencia de los psiquiatras. De manera similar, la función de la danza de la lluvia puede ser reemplazada por alguna otra ceremonia grupal. Es interesante que Malinowski mismo invoque en otro contexto "el principio de las posibilidades limitadas establecido por Goldenweiser. Dada una necesidad cultural definida, los medios para satisfacerla son poco numerosos y por lo tanto los mecanismos culturales que se originan como respuesta a esta necesidad están determinados por límites estrechos".[20] Es obvio que este principio implica por lo menos una moderada liberación del concepto de que todo ítem cultural es funcionalmente indispensable. Aun así, todavía puede ser demasiado restrictivo. De cualquier modo, sociólogos como Parsons y Merton admiten la existencia de "equivalentes funcionales" para ciertos ítems culturales. En su análisis general del funcionalismo, Merton ha insistido para que el concepto de indispensabilidad funcional de los ítems culturales sea reemplazado explícitamente por la noción de "alternativas funcionales o equivalentes funcionales o sustitutos funcionales".[21] Casualmente, esta idea tiene un paralelo interesante en el "principio de las soluciones múltiples" para problemas de adaptación en la evolución. Este principio, defendido por los biólogos funcionalistas, afirma que para un problema funcional dado (como la percepción de la luz) existe en general una serie de soluciones posibles, muchas de las cuales son utilizadas

[19] B. Malinowski: "Anthropology", *obra citada*, pág. 136; y *Magic, Science and Religion, and Other Essays, obra citada*, pág. 90. (Obsérvese la pretensión explicativa implícita en el uso de las palabras "de ahí".)

[20] B. Malinowski: "Culture". *Encyclopaedia of the Social Sciences*, IV, Nueva York, The Macmillan Company, 1931, pág. 626.

[21] Merton, *obra citada*, pág. 34. Compárese también con T. Parsons: *Essays in Sociological Theory, Pure and Applied*. Nueva York: The Free Press, 1949, pág. 58. Un intento interesante de establecer la existencia de alternativas funcionales en un caso específico es el de R. D. Schwartz: "Functional alternatives to inequality". *American Sociological Review*, 20, 1955, págs. 424-30.

por grupos de organismos diferentes, pero a menudo estrechamente ligados.[22]

Se debe destacar aquí que, en cualquier caso de análisis funcional, la cuestión de si hay equivalentes funcionales para un ítem dado *i* tiene significado concreto sólo si las condiciones internas y externas *c* en (4.1) se especifican claramente. De lo contrario, cualquier propuesta alternativa a *i*, digamos *i'*, podría no tener el status de equivalente funcional, dado que siendo diferente de *i*, el ítem *i'* tendría ciertos efectos sobre el estado interno y el ambiente de *s* que no los produciría *i*. Entonces, si se produjera *i'* en lugar de *i*, *s* no funcionaría en la misma situación interna y externa. Supongamos, por ejemplo, que el sistema de magia de un grupo primitivo cualquiera fuera reemplazado por una ampliación de su tecnología racional, además de algunas modificaciones en su religión, y que el grupo siguiera funcionando. ¿Se confirmaría así la existencia de un equivalente funcional al sistema original de magia? Se podría sostener una respuesta negativa basada en que, como resultado de la adopción de la pauta modíficada, el grupo hubiese cambiado tan profundamente respecto de algunas de sus características básicas (es decir, su estado interno, caracterizado por c_i, se hubiese modificado tan intensamente) que ya no se tratara del tipo original de grupo primitivo y que no habría ningún equivalente funcional de la magia que dejase intactos los caracteres "esenciales" del grupo. El uso coherente de este tipo de razonamiento pondría a salvo el postulado de la indispensabilidad de todo ítem cultural frente a cualquier desautorización empírica concebible... a costa de transformar la hipótesis empírica en una tautología encubierta.

Este proceso poco alentador debe ser evitado cuidadosamente. Pero, ¿qué puede establecer el análisis funcional a la manera general de (4.1) si la posibilidad de equivalentes funcionales de *i* no se descarta por definición? [23] Sea *I* la clase de todos aquellos ítems que son empíricamente suficientes para *n* bajo las circunstancias indicadas en (4.1), de modo tal que un ítem *j* se incluirá en *I*, por si acaso su realizacón en el sistema *s*, bajo condicones del tipo *c*, fueran empíricamente suficientes para asegurar el cumplimiento del requisito *n*. (La calificación de "empíricamente" es para indicar que el cumplimiento de *n* por *j* debe ser cuestión de hecho empírico y no sólo de lógica pura. Esta condición excluye de *I* los ítems triviales, tal como *n* mismo.) La clase *I* será entonces una clase de equivalentes funcionales en el sentido que le adjudicamos más arriba. Reemplacemos ahora la premisa (c) en (4.1) por el enunciado que sigue:

(c') *I* es la clase de todas las condiciones empíricamente suficientes para llenar el requisito *n* en el contexto determinado por el sistema *s*, en el ámbito *c*.

[22] Véase G. C. Simpson: *The Meaning of Evolution*. New Haven, Yale University Press, 1949, pág. 164 y sigs., 190, 342-43; y G. G. Simpson, C. S. Pittendrigh, L. H. Tiffany: *Life*, Nueva York, Harcourt, Brace & World, Inc.

[23] (Agregado en 1964.) Las conclusiones de esta sección fueron revisadas para corregir un error de la versión original que me señaló el profesor John R. Gregg.

En el mejor de los casos, las premisas (a), (b) y (c′) nos permitirán inferir que:

(4.2) Algunos de los ítems incluidos en la clase *I* están presentes en el sistema *s* en el momento *t*.

Pero esta conclusión no brinda razones para esperar que suceda algún ítem particular de *I* en lugar de uno de sus equivalentes funcionales. Estrictamente hablando, aun la débil conclusión (4.2) está garantizada sólo sobre la base de una ulterior premisa que dice que la clase *I* no es vacía, es decir, que existe por lo menos un ítem cuya ocurrencia aseguraría, por ley, el cumplimiento de *n*.

De modo que el análisis funcional no responde, tal como un razonamiento deductivo, a la presencia de un ítem determinado *i* que debe explicar. Entonces, ¿sería tal vez más adecuado construirlo como un razonamiento inductivo que expusiera la ocurrencia de *i* como muy probable bajo las circunstancias descriptas en las premisas? ¿No sería posible, por ejemplo, agregar a las premisas de (4.1) un enunciado más, a los efectos de que el prerrequisito funcional *n* sólo pueda ser llenado por *i* y por unas pocas alternativas funcionales especificables? ¿No podrían estas premisas confirmar la presencia de *i* como altamente probable? Este camino es muy poco prometedor ya que en la mayoría de los casos concretos, si no en todos, sería imposible especificar con cierta precisión la gama de pautas de conducta, instituciones o costumbres alternativas que bastarían para llenar un prerrequisito funcional o necesidad dados. Aun cuando pudiera caracterizarse esa gama, no hay en vista ningún método satisfactorio para dividirlo en algún número finito de casos y asignarle una probabilidad a cada uno.

Supongamos, por ejemplo, que la noción general de Malinowski sobre la función de la magia sea correcta. Cuando tratemos de explicar el sistema de magia de un grupo dado, ¿cómo podremos determinar todos los distintos sistemas de magia y pautas culturales alternativas que satisfagan las mismas necesidades funcionales del grupo que el sistema de magia actualmente vigente? ¿Cómo podremos asignar probabilidades de ocurrencia a cada uno de estos potenciales equivalentes funcionales? Es evidente que no existe respuesta satisfactoria a estas preguntas y tampoco los propulsores del análisis funcional pretenden concretar sus explicaciones de esta manera tan problemática.

No parece brindar tampoco ninguna ayuda construir las leyes generales implícitas en los enunciados (b) y (c) de (4.1) con forma estadística más bien que estrictamente universal, es decir, que expresen relaciones que son muy probables pero que no se aplican siempre. Las premisas así obtenidas, igual no excluirían las alternativas funcionales de *i* (cada una de las cuales haría altamente probable el cumplimiento de *n*) y por lo tanto siempre nos quedaría la dificultad básica: tomadas en su conjunto, las premisas no justificarían la sola presencia de *i* como muy probable.

Resumiendo, la información que típicamente nos proporciona el análisis funcional de un ítem *i*, no brinda un fundamento adecuado deductivo ni inductivo para preferir a *i* sobre cualquiera de sus otras alternativas. La impresión de que el análisis funcional efectivamente proporciona este

fundamento, y por lo tanto explica la ocurrencia de *i*, se debe sin duda y por lo menos en parte al beneficio del conocimiento *post facto*: cuando tratamos de explicar un ítem *i*, ya sabemos presumiblemente que *i* ha ocurrido.

Sin embargo, como decíamos hace un momento, el análisis funcional se puede construir como explicación deductiva con un explanandum muy débil del modo siguiente:

(4.3)

- (a) En el momento *t*, el sistema *s* funciona adecuadamente en un ámbito de tipo *c*
- (b) *s* funcional adecuadamente en un ámbito de tipo *c*, sólo si se satisface el requisito *n*
- (c′) *I* es la clase de condiciones empíricamente suficientes para *n* en el contexto determinado por *s* y *c*; *I* no es vacío
- (d′) Alguno de los ítems incluidos en *I* se halla presente en *s*, en el momento *t*

Sin embargo, este tipo de inferencia es bastante trivial, excepto cuando tenemos información adicional sobre los ítems contenidos en la clase *I*. Supóngase, por ejemplo, que en el momento *t* un perro (sistema *s*) goza de buena salud en un tipo de ámbito "normal" *c* que excluye mecanismos tales como corazones, pulmones y riñones artificiales. Supóngase además que en un ámbito del tipo *c* el perro puede gozar de buena salud sólo si su sangre circula adecuadamente (condición *n*). Entonces el esquema (4.3) nos lleva sólo a la conclusión de que de una u otra manera la sangre circula normalmente en el perro en el momento *t* —un resultado en verdad muy poco esclarecedor—. Sin embargo, si además sabemos de qué modo la sangre circula en esas circunstancias y si sabemos, por ejemplo, que el único rasgo que asegura una circulación adecuada (el único ítem en la clase *I*) es un corazón que funcione bien, entonces podremos llegar a una conclusión mucho más específica: la de que en el momento *t* el perro tiene un corazón sano. Pero si explicitamos otra información aquí usada, expresándola como una premisa adicional, entonces nuestro razonamiento puede reformularse bajo la forma (4.1), excepto que la premisa (c) ha sido reemplazado por el enunciado de que *i* es el *único* rasgo por el cual se puede satisfaces a *n* en el ámbito *c*. Como señaláramos más arriba, la conclusión (d) de (4.1) es la que se desprende en este caso.

Sin embargo, en general la información adicional de este tipo no se halla disponible y el alcance explicativo del análisis funcional se limita, pues, al precario rol esquematizado en (4.3).

5. *El alcance predictivo del análisis funcional*

Hemos hablado más arriba del significado predictivo de la explicación nomológica. Ahora nos preguntaremos si el análisis funcional puede tener uso predictivo.

En primer lugar, la discusión precedente muestra que la información típicamente proporcionada por el análisis funcional brinda, en el mejor de los casos, premisas que tienen las formas (a), (b), (c) en (4.1) y éstas

no son una base adecuada para la predicción deductiva o inductiva de un enunciado con forma (d) en (4.1). Por ello, el análisis funcional nos permite tan poco predecir, como explicar la ocurrencia de un ítem particular por el cual se pueda satisfacer un requerimiento funcional dado.

En segundo lugar, hasta el esquema explicativo (4.3), que es mucho menos ambicoso, carece de gran valor predictivo, ya que la derivación de la conclusión débil (e) se basa en la premisa (a). Si queremos inferir (e) con respecto a cierto momento futuro *t*, esa premisa no existe, ya que no sabemos si *s* funcionará adecuadamente en ese instante, o no. Por ejemplo: considérese una persona que manifiesta una ansiedad grave en forma creciente y supóngase que una condición necesaria para su funcionamiento adecuado sea que se la reprima con síntomas neuróticos o se la supere por algún otro medio. ¿Podremos predecir que se producirá realmente alguna de las formas de "adaptación" en la clase *I* así caracterizada? Es evidente que no, porque no sabemos si la persona en cuestión seguirá funcionando adecuadamente o si sufrirá algún colapso más o menos grave, tal vez hasta el punto de destruirse.

Es interesante destacar que también existe una limitación algo similar para el uso predictivo de las explicaciones nomológicas, aun en las ramas más avanzadas de la ciencia. Si tenemos que predecir por medio de las leyes de la mecánica clásica, por ejemplo, el estado en que se encontrará un sistema mecánico dado en un momento futuro *t*, no será suficiente saber el estado del sistema en un momento anterior t_0, como puede ser el presente. También necesitamos información acerca de las condiciones límite durante el intervalo de tiempo desde t_0 a *t*, es decir sobre las influencias externas que afectan al sistema en ese lapso. De un modo similar, la "predicción" en el primer ejemplo de que el nivel de agua quedaría igual a medida que el hielo se derritiera, supone que la temperatura del aire circundante será constante, digamos, y que no habrá influencias perturbadoras, tales como un terremoto o una persona que vuelque el recipiente. Cuando predecimos que un objeto que cae desde el techo del edificio Empire State llegará al suelo unos 8 segundos más tarde, suponemos que durante el lapso de su caída no se le aplica al objeto ninguna otra fuerza que la atracción gravitatoria de la Tierra. Así, en una formulación explícita y completa las predicciones nomológicas de este tipo deberían incluir, entre sus premisas, enunciados que especifiquen las condiciones límite obtenidas desde t_0 hasta el momento *t* al cual se refiere la predicción. Esto nos demuestra que aun las leyes y teorías de las ciencias físicas no nos permiten predecir en realidad ciertos aspectos del futuro sobre la base exclusiva de determinados aspectos del presente. La predicción también exige ciertos supuestos sobre el futuro. Pero en muchos casos de predicción nomológica existen buenos motivos inductivos disponibles en t_0 para suponer que durante el intervalo en cuestión el sistema estudiado quedará prácticamente "cerrado", es decir, libre de interferencias exteriores importantes (una ilustración de este caso sería, por ejemplo, la predicción de eclipses) o que las condiciones límite serán de un tipo especificado: situación ilustrada en predicciones de acontecimientos que ocurren bajo condiciones experimentalmente controladas.

El uso predictivo de (4.3) requiere a su vez una premisa concerniente al futuro, es decir (a). A menudo existe una incertidumbre considerable acerca de si (a) realmente será verdadera. Además, si en alguna instancia particular existieran buenas razones inductivas para considerar a (a) como verdadera, la predicción suministrada por (4.3) todavía sería bastante débil. Entonces el razonamiento lleva de la suposición inductivamente sustentada de que el sistema funcionará de manera adecuada en *t*, a la "predicción" de que cierta condición *n*, que es necesaria para ese funcionamiento, será satisfecha en *t* de uno u otro modo.

La necesidad de incluir supuestos acerca del futuro entre las premisas de los razonamientos predictivos puede evitarse tanto en las predicciones nomológicas como en las que están basadas en el análisis funcional, si nos quedamos satisfechos con conclusiones predictivas cuyo carácter no es categórico, sino sólo condicional o hipotético. Por ejemplo, (4.3) puede ser reemplazado por el siguiente razonamiento en el cual se evita la premisa (a) a costa de condicionalizar la conclusión:

(5.1)
(b) El sistema *s* funciona adecuadamente en un ámbito de tipo *c*, sólo si se cumple la condición *n*
(c′) *I* es la clase de condiciones empíricamente suficientes para *n* en el contexto determinado por *s* y *c*; *I* no es vacío
(d") Si *s* funciona adecuadamente en un ámbito de tipo *c* en el momento *t*, entonces alguno de los ítems en la clase *I* se halla presente en *s*, en el momento *t*

Esta posibilidad merece mencionarse porque parece que al menos algunas de las pretensiones de los que abogan por el análisis funcional pueden interpretarse como sosteniendo que nada más que el análisis funcional permite estas predicciones condicionales. Esta puede ser la intención de Malinowski, por ejemplo, cuando dice: "Si dicho análisis (funcional) revela que, tomando una cultura particular como un todo coherente podemos establecer una serie de determinantes generales a los cuales debe conformarse, podremos adelantar una serie de enunciados predictivos como guías para el trabajo de campo, como criterios para el tratamiento comparativo y como simples medidas en el proceso de adaptación cultural y cambio".[24] Los enunciados que especifiquen los determinantes en cuestión, presumiblemente tomarían la forma de premisas del tipo (b) y los "enunciados predictivos" tendrían entonces carácter hipotético.

Sin embargo, muchas de las predicciones y generalizaciones realizadas en el contexto del análisis funcional eluden esta forma condicional. Parten desde un enunciado de prerrequisito funcional o necesidad a la aseveración categórica del advenimiento de algún rasgo, institución u otro ítem, presumiblemente para llenar el requisito en cuestión. Consideremos por ejemplo la explicación funcional de Sait sobre el surgimiento del caudillo político: "El liderazgo es necesario y *puesto que* no se desarrolla fácilmente dentro del marco constitucional, el caudillo lo proporciona des-

[24] B. Malinowski: *A Scientific Theory of Culture, and Other Essays, obra citada*, pág. 38.

de fuera en forma cruda e irresponsable".[25] Tomemos la caracterización de Merton sobre una función de la maquinaria política; refiriéndose a varias maneras específicas en que el aparato político puede servir a los intereses empresarios, concluye: "Estas 'necesidades' del empresariado, tal como está constituido, no son satisfechas por las estructuras sociales convencionales y culturalmente aprobadas. *En consecuencia,* la organización extralegal pero más o menos eficiente del aparato político le provee estos servicios".[26] Cada uno de estos razonamientos, que son bastante típicos en el enfoque funcionalista, es una inferencia a partir de la existencia de determinado prerrequisito funcional a una aseveración categórica en que dicho prerrequisito será satisfecho de algún modo. ¿Cuál es el fundamento para las inferencias sugeridas por las palabras 'puesto que' y 'en consecuencia' en los párrafos recién mencionados? Cuando decimos *puesto que* se puso el cubo de hielo en agua caliente, se derritió; o bien que se hizo pasar la corriente y *en consecuencia* el amperímetro del circuito respondió, ambas inferencias pueden explicarse y justificarse refiriéndolas a ciertas leyes generales de las cuales estos casos particulares no son más que simples instancias y la lógica de las inferencias puede exponerse dándoles la forma del esquema (2.1). De manera similar, cada uno de los razonamientos funcionales que consideramos parecen presuponer claramente una ley general que establece que, dentro de ciertos límites de tolerancia o adaptabilidad, un sistema del tipo analizado cumplirá —sea de modo invariable o altamente probable— mediante el desarrollo de rasgos adecuados, los diversos requisitos funcionales (condiciones necesarias para su continuo funcionamiento adecuado) que pueden surgir a raíz de cambios en su estado interno o en su medio. Cualquier afirmación de este tipo, sea de forma estrictamente universal o estadística, será denominada *hipótesis* (*general*) *de autorregulación.*

Hasta tanto el análisis funcional del tipo que acabamos de ilustrar no se construya de modo que proponga o invoque implícitamente hipótesis adecuadas de autorregulación, no quedará claro qué conexiones indicarían las expresiones "puesto que", "en consecuencia" y otras del mismo estilo y cómo se establecerá objetivamente la existencia de esas relaciones.

Recíprocamente, si se formula una hipótesis precisa de autorregulación para sistemas de un tipo específico, será posible explicar y predecir categóricamente el cumplimiento de ciertos requisitos funcionales sobre la base de una simple información acerca de las necesidades antecedentes; entonces la hipótesis podrá ser probada objetivamente por medio de la corroboración empírica de sus prediccesiones. Tómese por ejemplo la afirmación de que si se corta en varios segmentos una hidra, cada uno de ellos crecerá para convertirse nuevamente en el animal completo. Este enunciado puede considerarse como una hipótesis acerca de un tipo específico de autorregulación en una determinada clase de sistema biológico. Su utilización con fines explicativos y predictivos es evidente y el éxito de sus predicciones lo confirma en gran medida.

[25] E. M. Sait: "Machine, Political", *Encyclopedia of the Social Sciences,* IX, Nueva York, The Macmillan Company, 1933, pág. 659. (La bastardilla es mía.)
[26] Merton, *obra citada,* pág. 76. (La bastardilla es mía.)

Vemos pues que siempre que el análisis funcional debe servir como base para la predicción categórica o para generalizaciones del tipo de las enunciadas por Sait y Merton, es de crucial importancia establecer hipótesis adecuadas de autorregulación en una forma objetivamente comprobable.

La bibliografía funcionalista incluye algunas generalizaciones formuladas explícitamente, del tipo de las que hemos referido aquí. Merton, por ejemplo, después de citar el párrafo de Sait mencionado más arriba, hace el siguiente comentario: "Puesto en términos más generalizados, *las deficiencias funcionales de la estructura oficial generan una estructura alternativa (no oficial) para llenar las necesidades existentes con algo más de eficacia*".[27] Este enunciado parece dirigido con claridad a explicar una hipótesis de autorregulación que subyace al análisis específico de Sait y que proporciona la racionalidad de su "puesto que". Radcliffe-Brown sugiere otra hipótesis de este tipo: "podría ser que dijéramos ... una sociedad que se encuentra en estado de desunión o incoherencia funcional ... no morirá, excepto en casos tan raros como una tribu australiana que fue arrollada por la fuerza destructiva del hombre blanco, sino que continuará luchando hacia ... algún tipo de salud social ..."[28]

Tal como lo sugiriera someramente más arriba, una formulación que se propone como hipótesis de autorregulación puede servir como base para la explicación o predicción, sólo si es lo suficientemente definida como para permitir la prueba empírica objetiva. Muchos de los principales representantes del análisis funcional han expresado su interés por desarrollar hipótesis y teorías que cumplan este requisito. Malinowski, por ejemplo, en un ensayo sugestivo titulado *Una teoría científica de la cultura* insiste en que "cada teoría científica debe partir y conducir a la observación. Debe ser inductiva y testable por la experiencia. En otras palabras, debe referirse a experiencias humanas que puedan definirse, que sean públicas, es decir, accesibles a cualquier observador y recurrentes; por lo tanto, plena de generalizaciones inductivas, o sea predictiva".[29] De manera similar, Murray y Kluckhohn dicen lo siguiente sobre el objetivo básico de su teoría funcional y además sobre cualquier otra "formulación" científica de la personalidad: "Los propósitos generales de la formulación son tres: (1) *explicar* los acontecimientos pasados y presentes; (2) *predecir* los acontecimientos futuros (cuyas condiciones se especifican); y (3) servir, si es necesario, como base para la selección de medidas efectivas de *control*".[30]

Sin embargo, desafortunadamente las formulaciones brindadas en el contexto de análisis funcionales concretos a menudo no cumplen estas premisas generales. Entre las muchas maneras en que se violan estas con-

[27] Merton, *obra citada*, pág. 73. (La bastardilla es del autor.)
[28] Radcliffe-Brown, *obra citada*, pág. 183.
[29] B. Malinowski: *A Scientific Theory of Culture, and Other Essays*, *obra citada*, pág. 67.
[30] Henry A. Murray y Clyde Kluckhohn: "Outline of a Conception of Personality". En Clyde Kluckhohn y Henry A. Murray (comps.): *Personality in Nature, Society, and Culture.* Nueva York, Knopf, 1950, págs. 3-32; cita de la pág. 7. (La bastardilla es del autor.)

diciones hay dos que deben ser consideradas muy especialmente dada su amplia difusión e importancia central en el análisis funcional. Las denominaremos (a) *especificación inadecuada del alcance* y (b) *uso no empírico de los términos funcionalistas clave* (como "necesidad", "requisito funcional", "adaptación" y otros). Consideraremos ambos defectos separadamente: el primero en lo que sigue de esta sección y el último en la siguiente.

La especificación inadecuada del alcance consiste en no indicar claramente el tipo de sistema al cual se refiere la hipótesis, o bien la gama de situaciones (los límites de tolerancia) dentro de las cuales se dice que esos sistemas desarrollan rasgos que llenarán los requisitos funcionales. La formulación de Merton, por ejemplo, no especifica la clase de sistemas sociales y de situaciones a los cuales se aplicarán las generalizaciones propuestas. Por lo tanto, tal como está enunciada no puede ser sometida a prueba empírica ni dársele uso predictivo.

La generalización tentativa expuesta por Radcliffe-Brown tiene una falla similar. Ostensiblemente se refiere a cualquier sociedad, pero las condiciones en que se pretende la supervivencia social están condicionadas por una cláusula sumamente indefinida ("excepto") que excluye la posibilidad de una testabilidad bien recortada. Hasta podría usarse la cláusula para proteger la generalización propuesta contra cualquier desconfirmación concebible. Si algún grupo social "muriera", este mismo hecho demostraría que las fuerzas destructoras eran tan abrumadoras como en el caso de la tribu australiana mencionada por Radcliffe-Brown. El uso sistemático de esta estrategia metodológica transformaría la hipótesis, naturalmente, en una tautología encubiera. Eso aseguraría su verdad, pero a costa de privarla de su contenido empírico: así construida la hipótesis no puede arrojar ninguna explicación o predicción.

Un comentario similar se aplica a la siguiente declaración de Malinowski en la cual la cláusula calificativa dudosa está en bastardilla: "Cuando consideramos una cultura que *no está a punto de ser destruida o de una ruptura total, sino que se trata de un ente cuyo funcionamiento es normal,* vemos que la necesidad y la respuesta están directamente relacionadas y adaptadas entre sí".[31]

Es cierto que las formulaciones de Radcliffe-Brown y Malinowski no necesariamente *tienen* que construirse como tautologías encubiertas y sin duda sus autores pretendían que fuesen afirmaciones empíricas. Pero en este caso, la vaguedad de las cláusulas calificativas las priva de la jerarquía de hipótesis empíricas definidas que podrían usarse con fines de explicación o predicción.

6. *Alcance empírico de los términos e hipótesis funcionalistas*

Un segundo defecto que puede viciar el rol científico de una hipótesis de autorregulación propuesta, consiste en usar términos del análisis fun-

[31] Malinowski: *A Scientific Theory of Culture, and Other Essays, obra citada,* pág. 94.

cional tales como "necesidad" y "funcionamiento adecuado (correcto)" [32] de una manera no empírica, es decir, sin darles una clara "definición operacional" o, dicho de un modo más general, sin especificar los criterios objetivos para su aplicación.[33] Si los términos funcionalistas se usan de esta manera, entonces las oraciones que los contienen no tienen un significado empírico claro: no conducen a predicciones específicas y por lo tanto no se prestan a la comprobación objetiva; ni tampoco, por supuesto, se los podrá utilizar con fines explicativos.

Es tanto más importante considerar aquí este punto cuanto que los términos claves del funcionalismo no sólo aparecen en hipótesis de autorregulación, sino también en otras oraciones funcionalistas de diversas calses, tales como las de los tipos (a), (b) y (d″) en (4.1), (4.3) y (5.1). El uso no empírico de términos funcionalistas puede, por lo tanto, excluir oraciones de distintos tipos de la jerarquía de hipótesis científica. Veamos algunos ejemplos.

Considérense en primer lugar los términos "prerrequisito funcional" y "necesidad" que se usan más o menos indistintamente en la literatura funcionalista, y que sirven para definir la palabra "función". "Enclavada en todo análisis funcional existe alguna concepción, tácita o explicitada, de los requerimientos funcionales del sistema bajo observación".[34] Efectivamente, "una definición (de función) es mostrar que las instituciones humanas, como también las actividades parciales que en ellas se desarrollan, están relacionadas a necesidades primarias (o sea biológicas) y derivadas (o sea culturales). Por lo tanto, la función significa siempre la satisfacción de una necesidad..." [35]

¿Cómo se define este concepto de necesidad? Malinowski nos brinda una respuesta explícita. "Entiendo, pues, por necesidad, el sistema de condiciones en el organismo humano, en el ámbito cultural y en la relación de ambos con el medio, que son suficientes y necesarias para la supervivencia del grupo y del organismo".[36] Esta definición parece clara y directa; no obstante, ni siquiera está muy de acuerdo con el uso que el propio Malinowski da al concepto de necesidad ya que, muy plausiblemente, dis-

[32] De acuerdo con la práctica aceptada ampliamente en la lógica contemporánea, entenderemos por términos cierto tipo de palabras u otras expresiones lingüísticas, agregando que un término expresa o significa un concepto. Por ejemplo, diremos que el término "necesidad" significa el concepto de necesidad. Como lo demuestra este ejemplo, nos referimos o mencionamos una expresión lingüística, utilizando una denominación que consiste simplemente en encerrar la expresión entre comillas.

[33] Una discusión general sobre la naturaleza y significado de los criterios "operacionales" y aplicación para los términos usados en la ciencia empírica y otras referencias a la bioliografía sobre el tema pueden encontrarse en C. G. Hempel: *Fundamentals of Concept Formation in Empirical Science*. University of Chicago Press, 1952, secciones 5-8; también en las presentaciones del simposio sobre estado actual del operacionalismo por G. Bergmann, P. W. Bridgman, A. Grunbaum, C. G. Hempel, R. B. Lindsay, H. Margenau y R. J. Seeger que forman el capítulo II de Philipp G. Frank (comp.): *The Validation of Scientific Theories*. Boston, The Beacon Press, 1956.

[34] Merton, *obra citada*, pág. 52.

[35] Malinowski: *A Scientific Culture, and Other Essays, obra citada*, pág. 159.

[36] Malinowski: *ibíd.*, pág. 90.

tingue una considerable cantidad de necesidades diferentes que divide en dos grandes grupos: necesidades biológicas primarias y culturales derivadas. Estas últimas incluyen las necesidades "tecnológicas, económicas, legales y hasta mágicas, religiosas o éticas".[37] Pero si cada una de estas necesidades no sólo representa una condición necesaria sino también suficiente para la supervivencia, entonces es evidente que la satisfacción de una necesidad única bastaría para asegurar la supervivencia y las otras necesidades no constituirían de ningún modo condiciones necesarias para lograrla. Por lo tanto, parecería razonable suponer que lo que Malinowski se proponía era construir las necesidades de un grupo como un conjunto de condiciones que son necesarias por separado y suficientes en conjunto para su supervivencia.[38]

Sin embargo, aunque se corrija este defecto lógico secundario, no se remedia uno más grave en la definición de Malinowski, que reside en el aspecto engañoso de claridad de la frase "supervivencia del grupo y del organismo". Con respecto al organismo biológico, la palabra "supervivencia" tiene un significado bastante claro, si bien aquí también es necesaria una mayor clarificación. Cuando hablamos de necesidades o requerimientos biológicos (por ejemplo los requisitos diarios mínimos para adultos de los distintos minerales y vitaminas) los construimos no como condiciones para la mera supervivencia, sino como condiciones de persistencia dentro de un estado "normal" o "sano" o de retorno a él, o bien un estado en el cual el sistema es un "todo que funciona correctamente". Para poner a prueba la comprobación objetiva de las hipótesis funcionales es esencial, pues, que las definiciones de necesidades o prerrequisitos funcionales se complementen con criterios razonablemente claros y objetivamente aplicables de lo que debe considerarse un estado sano o un funcionamiento normal de los sistemas en consideración.

Los análisis funcionales en psicología, sociología y antropología tienen una necesidad aun más urgente de criterios empíricos objetivos del tipo al que nos hemos referido, porque la caracterización de necesidades que son condiciones necesarias para la supervivencia psicológica y emocional del individuo o de un grupo es lo suficientemente vaga como para permitir y hasta incitar a interpretaciones subjetivas muy distintas.

Algunos autores caracterizan el concepto de prerrequisito funcional o el de función sin utilizar la palabra "supervivencia", que tiene una engañosa apariencia de claridad. Merton, por ejemplo, afirma: "*Funciones* son todas aquellas consecuencias observadas que ayudan a la adaptación o

[37] Malinowski, *ibíd.*, pág. 172; véase también *ibíd.*, pág. 91 y sigs.

[38] En algunas de sus disquisiciones, Malinowski también descarta, por implicación, el concepto de función entendida como satisfacción de un estado que por lo menos es *necesario* para la supervivencia del grupo o del organismo. Por ejemplo, en el trabajo que incluye los dos párrafos recién citados en el texto, Malinowski comenta sobre la función de algunos logros culturales complejos lo siguiente: "Tomemos el avión, el submarino o la locomotora. Es obvio que el hombre no tiene necesidad de volar, ni de hacerle compañía a los peces y moverse en un medio para el cual no está ni adaptado anatómicamente ni preparado fisiológicamente. Por lo tanto, para definir la función de cualquiera de estos inventos no podemos afirmar el verdadero motivo de su aparición en función de una necesidad metafísica". *Ibíd.*, págs. 118-119.

ajuste de un sistema dado. *Disfunciones* son aquellas consecuencias observadas que disminuyen la adaptación o ajuste del sistema".[39] Radcliffe-Brown caracteriza la función de un ítem como su contribución al mantenimiento de una cierta unidad en un sistema social "que podemos llamar unidad funcional. La que podemos definir como una condición en la que todas las partes del sistema social colaboran con un grado suficiente de armonía y coherencia interna, es decir, sin producir conflictos persistentes que no puedan resolverse ni regularse".[40] Al igual que las definiciones en que entra la supervivencia, estas caracterizaciones alternativas, si bien son sugestivas, distan mucho de brindar significados empíricos claros para los términos clave del análisis funcional. Por ejemplo, los conceptos de ajuste y adaptación requieren la especificación de alguna norma; en caso contrario no tienen significado definido y corren el peligro de ser usados tautológicamente, o bien con tintes valorativos subjetivos.

El uso tautológico puede basarse en la interpretación de que *cualquier* respuesta de un sistema dado constituye un ajuste, en cuyo caso la noción de que todo sistema se puede ajustar a algún conjunto de circunstancias se convierte en una verdad trivial. Algunas instancias del análisis funcional se acercan peligrosamente a este procedimiento, como lo demuestra la siguiente afirmación: "Tenemos así una explicación del suicidio y de numerosos efectos aparentemente antibiológicos que serían otras tantas formas para aliviar un sufrimiento intolerable. El suicidio no tiene valor *de adaptación* (sobrevivencia) pero sí *de ajuste* para el organismo. El suicidio es *funcional* porque elimina la tensión dolorosa".[41]

Considérese la formulación de Merton de uno de los supuestos del análisis funcional: "...cuando *el balance neto del conjunto de consecuencias* de una estructura social existente es netamente disfuncional, se desarrolla una fuerte e insistente presión para que cambie".[42] No habiendo criterios empíricos claros de adaptación (y por lo tanto de disfunción) es posible encarar esta formulación como una tautología encubierta y hacerla impenetrable a la desconfirmación empírica. Merton es consciente del peligro; en otro contexto comenta que la noción de requisitos funcionales de un sistema dado "sigue siendo uno de los conceptos más nebulosos y empíricamente cuestionables de la teoría funcional. Tal como lo utilizan los sociólogos, el concepto de requerimiento funcional tiende a ser tautológico o *ex post facto*".[43] Otros autores, tales como Malinowski[44] y Parsons[45] han prevenido contra el uso tautológico y las generalizaciones *ad hoc* de los prerrequisitos funcionales.

Al faltar criterios empíricos de ajuste o adaptación, también se corre

[39] Merton, *obra citada*, pág. 51. (La bastardilla es del autor.)
[40] Radcliffe-Brown, *obra citada*, pág. 181.
[41] Murray y Kluckhohn, *obra citada*, pág. 15. (La bastardilla es de los autores.)
[42] Merton, *obra citada*, pág. 40.
[43] Merton, *obra citada*, pág. 52.
[44] Véase, por ejemplo, Malinowski: *A Scientific Theory of Culture, and Other Essays, obra citada*, págs. 169-170, pero compárese también con las págs. 118-119 de la misma obra.
[45] Véase, por ejemplo, T. Parsons: *The Social System.* Nueva York, The Free Press, 1951, 29, n. 4.

el peligro de que cada investigador proyecte sobre este concepto (y por lo tanto también sobre el concepto de función) sus propias normas éticas de lo que constituiría un ajuste "adecuado" o "correcto" de un sistema dado —lo cual ha sido señalado muy claramente por Levy—.[46] Este procedimiento, obviamente, privaría a las hipótesis funcionalistas de su jerarquía de afirmaciones científicas precisas, y objetivamente comprobables. Como lo destaca Merton: "Si queremos que la teoría sea productiva, debe ser lo suficientemente *precisa* como para *determinar*. La precisión es un elemento integral del cirterio de *testabilidad*".[47]

Es, entonces, esencial para que el análisis funcional sea un procedimiento científico, que sus conceptos claves se erijan explícitamente con respecto a cierta norma de supervivencia o ajuste. Esta norma debe especificarse para cada análisis funcional y en general variará de caso en caso. En el estudio funcional de un sistema dado *s*, la norma estaría indicada por la especificación de alguna clase o gama R de estados posibles *s*, entendiéndose que se considerará a *s* como "supervivencia con buen funcionamiento" o "adaptación correcta en condiciones cambiantes" en el caso de que *s* siga o retorne, después de alguna perturbación, a un estado dentro de los límites de R. Una necesidad o requisito funcional del sistema *s* con respecto a R es, pues, una condición necesaria para que el sistema permanezca o retorne al estado en R. Asimismo, la función referente a R de un ítem *i* en *s* consiste en que *i* produzca la satisfacción de tal requisito funcional.

En el campo de la biología, el análisis de Sommerhoff sobre adaptación, aptitud y otros conceptos relacionados constituye una excelente ilustración de un estudio formal en el que la relativización de los conceptos funcionalistas centrales se ha hecho totalmente explícita.[48] La necesidad de una relativización como ésta también ha sido aclarada por Nagel, quien señala que "pretender que un cambio dado sea funcional o disfuncional debe entenderse como relacionado con una G (o conjunto G)"[40] especificada, donde las G son características cuya preservación sirve como norma para definir el ajuste o supervivencia. En sociología, el análisis de Levy sobre estructura de la sociedad[50] construye con toda claridad los conceptos funcionalistas clave en un sentido relativo como el que antes esbozamos.

Sólo si los conceptos clave del análisis funcional se relativizan de esta

[46] Marion J. Levy (h.): *The Structure of Society*. Princeton, Princeton University Press, 1952, pág. 76 y sigs.

[47] R. K. Merton: "The Bearing of Sociological Theory on Empirical Research". En Merton: *Social Theory and Social Structure, obra citada*, págs. 85-101; cita de pág. 98. (La bastardilla es del autor).

[48] Véase G. Sommerhoff: *Analytical Biology*. Nueva York, Oxford University Press, 1950.

[49] E. Nagel: "A Formalization of Functionalism", *obra citada*, pág. 269. Véase también el último párrafo del mismo artículo (págs. 282-283).

[50] Levy habla de eufunción y disfunción de una unidad (es decir de un sistema) y caracteriza estos conceptos referidos a la "unidad tal como se la define". Señala que la relativización es necesaria "porque se debe volver a la definición de unidad para determinar si se está produciendo o no la 'adaptación o ajuste' que permite la persistencia o falta de persistencia de esa unidad". Levy, *ibíd.*, págs. 77-78.

manera, las hipótesis que los incluyen podrán tener la jerarquía de supuestos o aseveraciones determinadas y objetivamente comprobables. Sólo en ese caso esas hipótesis pueden integrar significativamente razonamientos tales como los descriptos en (4.1), (4.3) y (5.1).

Si bien esta relativización puede otorgar contenido empírico definido a las hipótesis funcionalistas que sirven de premisas o conclusiones a los razonamientos, el alcance explicativo y predictivo de estos últimos sigue siendo tan limitado como lo vimos en las secciones 4 y 5, ya que nuestro veredicto sobre la fuerza lógica de aquellos razonamientos dependía únicamente de su estructura formal y no del significado de sus premisas y conclusiones.

Por lo tanto, sigue siendo cierto, aun para una versión correctamente relativizada del análisis funcional, que su poder explicativo es bastante limitado. En especial, no proporciona una explicación acerca de por qué un ítem *i* ocurre en el sistema *s* en lugar de alguno de sus equivalentes funcionales. El significado predictivo del análisis funcional es prácticamente nulo, excepto en aquellos casos en que pueden establecerse hipótesis adecuadas de autorregulación. Esta determina que dentro de una gama especificada C de circunstancias, un sistema dado *s* (o bien un sistema de cierto tipo S del cual *s* es una instancia) se autorregula con respecto a una gama especificada R de estados, es decir, que después de una perturbación que desplaza a *s* a un estado fuera de R, pero que no cambia las circunstancias internas y externas de *s* más allá de la gama especificada C, el sistema *s* volverá a su estado en R. Un sistema que cumple una hipótesis de este estilo puede decirse que se *autorregula con respecto a R.*

Los sistemas biológicos brindan muchos ejemplos de autorregulación. Por ejemplo, se mencionó antes la capacidad regenerativa de la hidra. Consideremos pues el caso en donde se secciona un segmento más o menos grande del animal y el resto crece hasta convertirse otra vez en una hidra completa. La clase R está constituida por el estado en que la hidra se completa. La caracterización de la gama C tendrá que incluir (a) una especificación acerca de la temperatura y composición química del agua en la cual la hidra cumplirá su proeza regenerativa (es evidente que no tendrá una sola composición sino unas cuantas diferentes: las concentraciones de distintas sales, por ejemplo, tendrán un valor dentro de una escala específica y tal vez limitada; lo mismo para la temperatura del agua) y (b) un enunciado acerca del tipo y tamaño del segmento que puede ser removido sin que ello impida la regeneración.

Sin duda, una de las tareas más importantes del análisis funcional en psicología y en las ciencias sociales será la de determinar hasta qué punto pueden encontrarse estos fenómenos de autorregulación y cómo podrán ser representados por sus correspondientes leyes.

7. *Análisis funcional y teleología*

Sean cuales fueren las leyes específicas que puedan descubrirse mediante la investigación que siga estos lineamientos, el tipo de explicación

y predicción que posibilitan no difiere, en cuanto a su carácter lógico, del de las ciencias físicas.

Es bien cierto que las hipótesis de autorregulación, que podrían ser resultado de una investigación funcionalista exitosa, exhiben un carácter teleológico, ya que afirman que dentro de condiciones especificadas los sistemas de cierto tipo tenderán a un estado en la clase R, que de este modo adquiere apariencia de causa final, determinando el comportamiento del sistema.

Ante todo, sería sencillamente insostenible decir de un sistema *s* autorregulado con respecto a R, que su futuro retorno a (un estado en) R es una "causa final" que determina su comportamiento actual. Porque aun cuando *s* se autorregula con respecto a R y ha sido desplazado a un estado fuera de R, la posibilidad futura de su retorno a R puede no producirse jamás: en el proceso de retorno a R, *s* puede verse expuesto a otras perturbaciones que lo arrojen fuera de la gama permitida C y que conduzcan a la destrucción de *s*. Por ejemplo, cuando a la hidra se le secciona un solo tentáculo, se producirán rápidamente ciertos procesos de regeneración, que no pueden explicarse de manera teleológica refiriéndolos a una causa final que consistiría en la eventualidad futura de que la hidra se completase otra vez. Esa eventualidad puede no producirse nunca ya que, durante el proceso de regeneración y antes de su compleción, el animal puede sufrir nuevos daños, irreparablemente graves y morir. Por lo tanto, lo que justifica los cambios actuales de un sistema autorregulado *s*, no es la "eventualidad futura" de que *s* esté en R, sino más bien la *disposición actual* de *s* de retornar a R. Es esta disposición la que está expresada en la hipótesis de autorregulación que gobierna al sistema *s*.

Cualquiera que sea el carácter teleológico que pueda atribuirse a la explicación o predicción funcionalista que invoque hipótesis (debidamente relativizadas) de autorregulación, éste reside sólo en la circunstancia de que tales hipótesis aseveran una tendencia de ciertos sistemas a mantenerse o retornar a ese tipo de estado. Leyes de esa índole que atribuyen un comportamiento característico orientado hacia una meta a sistemas de tipos específicos, no son ajenas a la física y a la química. Por el contrario, estos últimos campos son los que proveen los ejemplos más comprensibles de sistemas autorreguladores y leyes relacionadas. Así, el líquido de una vasija volverá a su estado de equilibrio en que la superficie es horizontal después de una perturbación mecánica; una banda elástica, después de ser estirada (dentro de ciertos límites) volverá a su forma original cuando se la suelta. Muchos sistemas controlados por mecanismos de realimentación negativa, como la locomotora provista de un regulador de velocidad, el torpedo con cabeza buscadora del blanco o el avión guiado por piloto automático, muestran, dentro de límites determinables, la autorregulación con respecto a alguna clase especial de estados.

En todos estos casos, las leyes de autorregulación que exhiben los sistemas en cuestión están sujetas a la explicación por subsunción a leyes generales que tienen una forma causal más obvia. Sin embargo, esto ni siquiera es esencial ya que las mismas leyes de autorregulación son causales, en el sentido amplio de que afirman que para sistemas de un tipo

específico, cualquier clase de "estados iniciales" diferentes (cualquiera de los estados permisibles de perturbación) llevará a algún tipo de estado final. Como lo habíamos demostrado en formulaciones anteriores, las hipótesis funiconalistas, incluyendo las de autorregulación, pueden ser expresadas sin utilizar para nada la terminología teleológica.[51]

No existen, pues, razones sistemáticas para atribuir al análisis funcional un carácter sui géneris que no se encuentre en las hipótesis y teorías de las ciencias naturales y en las explicaciones y predicciones que se basan en ellas. Sin embargo, desde el punto de vista psicológico la idea de función a menudo permanece estrechamente ligada a la de propósito, y algunas publicaciones funcionalistas han estimulado sin duda esta asociación usando una terminología que atribuye al comportamiento autorregulado de un sistema dado el carácter de una acción intencional. Por ejemplo, Freud, al hablar de la relación entre síntomas neuróticos y ansiedad, usa un lenguaje fuertemente teleológico cuando sostiene que "los síntomas se crean para apartar o librar al yo de la situación de peligro".[52] Las citas transcriptas en la sección 3 proporcionan más ejemplos ilustrativos. Merton, que rechaza esta práctica enfáticamente y en forma muy explícita, ha reunido algunos ejemplos interesantes de artículos sociológicos y antropológicos en los que se confunden los conceptos de función y de propósito.[53]

Parece verosímil que precisamente esta asociación psicológica del concepto de función con el de propósito (si bien de manera sistemática no tiene justificación) considere en gran medida lo sugestivo y la aparente plausibilidad del análisis funcional como modo de explicación. Así parecería que nos permite "comprender" los sistemas autorreguladores de todo tipo de función de propósitos o motivos, del mismo modo en que "comprendemos" nuestra propia conducta intencional y la de los demás. Una explicación que haga referencia a motivos u objetivos puede tener una legitimidad perfecta en el caso de la conducta intencional y de sus efectos. Una explicación de este tipo tendría carácter causal y detallaría entre los antecedentes causales de una acción determinada o de su resultado, ciertos propósitos o motivos atribuibles al agente, así como sus convicciones acerca de cuáles serán los mejores medios disponibles para lograr sus objetivos. Este tipo de información sobre propósitos y convicciones puede aún servir como punto de partida para explicar un rasgo autorregulador en un artefacto fabricado por el hombre. Por ejemplo, para explicar la presencia del regulador de velocidad en una locomotora, parecería bastante razonable referirse al destino que su inventor procuraba darle, sus convicciones con-

[51] Para encontrar interesantes debates sobre otros aspectos concernientes a la "explicación teleológica", especialmente en lo que se refiere a sistemas autorreguladores, véase R. B. Braithwaite: *Scientific Explanation*. Cambridge, Cambridge University Press, 1953, capítulo X; y E. Nagel: "Teleological Explanation and Teleological Systems". En S. Ratner (comp.): *Vision and Action: Essays in Honor of Horace Kallen on His Seventieth Birthday*, New Brunswick, N. J., Rutgers University Press, 1953; reproducido en H. Feigl y M. Brodbeck (comps.): *Readings in the Philosophy of Science*. Nueva York; Appleton-Century-Crofts, Inc., 1953.

[52] Freud, *obra citada*, pág. 112.

[53] Merton: "Manifest and Latent Functions", *obra citada*, págs. 23-25, 60 y siguientes.

cernientes a los hechos físicos y la tecnología de la cual disponía. Digamos que estas referencias podrían ofrecernos una explicación probabilística de la presencia del regulador pero no revelarían por qué funcionaba como mecanismo autorregulador de seguridad. Para explicar esto último tendríamos que explayarnos sobre la construcción de la máquina y las leyes de la física, no sobre las intenciones y convicciones de su inventor. (También se puede brindar una explicación basada en motivos y convicciones para determinados ítems que en la realidad no funcionan como era de esperarse; por ejemplo, algunas prácticas supersticiosas, las máquinas voladoras que no tuvieron éxito, las políticas económicas ineficaces, etc.) Más aún —y éste es el punto crucial en nuestro contexto— para la mayoría de los fenómenos autorreguladores que están bajo los alcances del análisis funcional, la atribución de propósitos es un desplazamiento ilegítimo del concepto de propósito desde su ámbito significativo de aplicación a uno mucho mayor, donde carece de justificación empírica objetiva. En el contexto de la conducta intencional o manifiesta de individuos o grupos existen varios métodos para comprobar si los motivos o propósitos supuestos están presentes verdaderamente en una situación dada. Una manera bastante directa sería la de entrevistar a los agentes en cuestión, pero existen muchos procedimientos "operacionales" alternativos de carácter más indirecto. Por lo tanto, las hipótesis explicativas en función de propósitos son susceptibles de pruebas razonablemente objetivas. No obstante, estos criterios empíricos están ausentes en otros casos de sistemas autorregulados y entonces no tendrá sentido científico atribuirles algún propósito. Igualmente tenderá a fomentar la ilusión de que se ha adquirido una comprensión profunda, que se gana percepción sobre la naturaleza de estos procesos asemejándolos a un tipo de conducta con la que estamos muy familiarizados por la experiencia cotidiana. Considérese por ejemplo la ley de "adaptación a un fin obvio" planteada por el sociólogo L. Gumplowicz, quien pretende que se aplica tanto al ámbito natural como al social. Para esto último afirma que "todo desarrollo social, toda entidad social, sirve a un fin definido, aun cuando su valor y moralidad puedan ser cuestionados. La ley universal de adaptación significa simplemente que ningún esfuerzo, ningún cambio de condiciones carece de propósito en cualquier ámbito de los fenómenos. De allí que deba aceptarse la racionalidad inherente de todos los hechos y condiciones sociales".[54] Se sugiere poderosamente que esta supuesta ley nos permite comprender la dinámica social dada su estrecha analogía con la conducta intencional dirigida hacia la consecución de un fin. Sin embargo, la ley carece totalmente de sentido empírico ya que no se ha dado ninguna interpretación empírica a términos tan claves como "fin", "falta de propósito" y "racionalidad inherente" para los contextos en los cuales se aplican. Esta "ley" no asevera nada y por lo tanto no puede de ningún modo explicar los fenómenos sociales (u otros).

El libro de Gumplowicz es varias décadas anterior a los escritos de Malinowski y de otros conspicuos funcionalistas. Estos autores contempo-

[54] L. Gumplowicz: *The Outlines of Sociology*, traducido por F. W. Moore. American Academy of Political and Social Science, 1899, págs. 79-80.

ráneos han sido más cautos y sutiles para plantear sus ideas. Así y todo, hay ciertas afirmaciones bastante importantes en la actual bibliografía funcionalista que ciertamente recuerdan la formulación de Gumplowicz, al sugerir una comprensión de los fenómenos funcionales a imagen y semejanza de la conducta manifiesta deliberada o de los sistemas que funcionan de acuerdo con un diseño preconcebido. Las opiniones siguientes ilustrarán este punto: "(La cultura) es un sistema de objetos, actividad y actitudes en la que cada parte existe como medio para un fin" [55] y "La perspectiva funcional de la cultura insiste pues sobre el principio de que en todo tipo de civilización, las costumbres, objeto material, idea o creencia cumplen alguna función vital, tienen una misión que cumplir, representan una parte indispensable dentro de un todo integrado".[56] Estos enunciados expresan lo que Merton denomina, en una discusión crítica, el postulado del funcionalismo universal.[57] Merton califica este postulado de prematuro.[58] El razonamiento presentado en la sección previa muestra que, faltando una interpretación empírica clara de los términos clave del funcionalismo, ese postulado es algo más que prematuro: es empíricamente vacío. Las formulaciones de este tipo pueden evocar una sensación de percepción y comprensión, equiparando los desarrollos socioculturales con la conducta intencional, reduciéndolos así a fenómenos que nos son muy conocidos. Pero la explicación y la comprensión científicas no son simples reducciones a lo que es conocido; si así fuera, la ciencia no se empeñaría en explicar fenómenos conocidos. Además, los progresos más importantes de nuestra comprensión científica del mundo se logran a menudo por medio de nuevas teorías, como la de los cuantos, que suponen tipos de objetos o procesos bastante poco conocidos que no pueden observarse directamente y que suelen exhibir características extrañas y hasta paradójicas en apariencia. Determinados fenómenos han sido científicamente comprendidos cuando se los puede encasillar en una teoría o sistema de leyes comprobables y adecuadamente confirmados. Las virtudes del análisis funcional tendrán que ser juzgadas por su capacidad de lograr ese tipo de comprensión.

8. *El papel heurístico del análisis funcional*

Las consideraciones precedentes sugieren que lo que a menudo ha recibido el nombre de "funcionalismo" debe encararse no tanto como un cuerpo de doctrinas o teorías que propongan principios tremendamente generales, tales como el principio del funcionalismo universal, sino más bien como un programa de investigación guiado por ciertas máximas heurísticas o "hipótesis de trabajo". La idea del funcionalismo universal, por ejemplo, que es insostenible cuando se la formula como ley empírica o

[55] Malinowski: *A Scientific Theory of Culture, and Other Essays, obra citada,* pág. 150.

[56] Malinowski: "Anthropology", *obra citada,* pág. 133.

[57] Merton: "Manifest and Latent Functions", *obra citada,* pág. 30 y sig.

[58] *Ibíd.,* pág. 31.

principio teórico totalizador, puede aplicarse con mayor beneficio si expresa una directiva para la investigación; es decir, que busque aspectos autorreguladores específicos del sistema social y otros, para examinar de qué manera las distintas características de un sistema pueden contribuir a ese modo particular de autorregulación. (De todos los "axiomas generales del funcionalismo" sugeridos por Malinowski —y que él consideraba demostrados por toda la evidencia empírica pertinente— puede ofrecerse una interpretación similar como máximas heurísticas para la investigación empírica.)[59]

En biología, por ejemplo, la contribución del enfoque funcionalista no está dada por la afirmación totalizadora de que todas las características de cualquier organismo satisfagan alguna necesidad y por lo tanto cumplan una función. En este nivel de generalización, la frase tiende a no tener sentido, o bien a ser disimuladamente tautológica o empíricamente falsa (según si el concepto de necesidad no recibe una interpretación empírica clara o si se lo trata de un modo autologizante o se le da una interpretación empírica específica). En cambio, los estudios funcionales en biología se han dedicado a demostrar, por ejemplo, cómo procesos homeostáticos y regenerativos específicos en diferentes especies contribuyen a mantener y desarrollar al organismo vivo. A continuación (a) han examinado con precisión cada vez mayor la naturaleza y los límites de dichos procesos (esto se refiere básicamente al establecimiento de distintas hipótesis empíricas específicas o leyes de autorregulación) y (b) han explorado los mecanismos fisiológicos y fisicoquímicos subyacentes, así como las leyes que los gobiernan, a fin de lograr una comprensión teórica más profunda de los fenómenos estudiados.[60] Existen corrientes similares en el estudio de los aspectos funcionales de los procesos psicológicos, incluyendo, por ejemplo, la formación de síntomas en la neurosis.[61]

El análisis funcional en psicología y en las ciencias sociales, como también en biología, puede concebirse, al menos idealmente, como un programa de investigación orientado a determinar los aspectos y grados en que se autorregulan distintos sistemas, en el sentido que le hemos dado aquí. Ese concepto está claramente reflejado en el artículo de Nagel "Formalización del funcionalismo"[62] que desarrolla un esquema analítico inspirado en el análisis formal de autorregulación en biología de Sommer-

[59] Malinowski: *A Scientific Theory of Culture, and Other Essays, obra citada,* pág. 150.

[60] Una descripción de este tipo de enfoque de los procesos homeostáticos en el cuerpo humano puede encontrarse en Walter B. Cannon: *The Wisdom of the Body.* Nueva York, W. W. Norton & Company, Inc., edición revisada 1939.

[61] Véase, por ejemplo, J. Dollard y N. E. Miller: *Personality and Psychotherapy.* Nueva York, McGraw-Hill Book Company, Inc., 1950, capítulo XI, "How Symptoms are Learned" y léase especialmente las págs. 165-66.

[62] E. Nagel: "A Formalization of Functionalism", *obra citada.* Véase también la discusión más general sobre el análisis funcional incluido en el trabajo de Nagel "Concept and Theory Formation in the Social Sciences" en *Science, Language, and Human Rights*; American Philosophical Association Eastern Division, Volume 1, Filadelfia, University of Pennsylvania Press, 1952, págs. 43-64. Reimpreso en J. L. Jarret y S. M. McMurrin (comps.): *Contemporary Philosophy,* Nueva York, Henry Holt & Co., Inc., 1954.

hoff [63] y muy similar al mismo, utilizándolo para mostrar y clarificar la estructura del análisis funcional, especialmente en sociología y antropología.

El enfoque funcionalista ha probado que puede ser esclarecedor, sugestivo y fructífero en muchos contextos. Si las ventajas que ofrece han de aprovecharse al máximo, parecería útil y también necesario, continuar la investigación de relaciones funcionales específicas hasta tal punto que puedan expresarse en términos de hipótesis razonablemente precisas y objetivamente comprobables. Inicialmente, al menos, estas hipótesis tendrán un alcance bastante limitado, lo cual las equipararía con la situación actual en biología, donde los tipos de autorregulación y uniformidades que exhiben, varían de especie en especie. Con el tiempo, estas "generalizaciones empíricas" de alcance limitado podrán proporcionar la base para una teoría más general de sistemas autorreguladores. Hasta qué punto se alcanzarán dichos objetivos, no puede decidirse *a priori* por medio del análisis lógico o de la reflexión filosófica: la respuesta se hallará en la investigación científica intensa y rigurosa.

[63] Sommerhoff, *obra citada.*

Capítulo XII

Aspectos de la explicación científica [1]

1. *Introducción*

Entre los muchos factores que han estimulado y sostenido la investigación en los diversos campos de la ciencia empírica, hay dos perdurables preocupaciones humanas que han suministrado el principal impulso a los esfuerzos científicos del hombre.

Uno de ellos es de naturaleza práctica. El hombre no sólo quiere sobrevivir en el mundo, sino también mejorar su posición estratégica dentro de él. Esto hace que sea importante poder hallar maneras confiables de prever cambios en su ambiente y, si es posible, controlarlos para usarlos en su propio provecho. La formulación de leyes y teorías que permiten la predicción de sucesos futuros se cuentan entre las más altas realizaciones de la ciencia empírica; y la medida en la cual ellas responden al anhelo del hombre de previsión y control la indica el vasto ámbito de sus aplicaciones prácticas que van desde las predicciones astronómicas hasta los pronósticos meteorológicos, demográficos y económicos, y desde la tecnología fisicoquímica y biológica hasta el control psicológico y social.

La segunda motivación básica de las indagaciones científicas del hombre es independiente de tales preocupaciones prácticas. Reside en su pura curiosidad intelectual, en su profundo y persistente deseo de conocer y de comprenderse a sí mismo y a su mundo. Tan intenso es este deseo, en verdad, que en ausencia de un conocimiento más confiable, a menudo se acude a los mitos para llenar el abismo. Pero al mismo tiempo, muchos de esos mitos ceden el terreno a concepciones científicas acerca del cómo y el porqué de los fenómenos empíricos.

[1] Este ensayo no ha sido impreso con anterioridad. Incluye, sin embargo, algunos pasajes de los siguientes artículos: "Deductive-Nomological vs. Statistical Explanation", en Herber Feigl y Grower Maxwell (comps.): *Minnesota Studies in the Philosophy of Science*, vol. III, Minneapolis, University of Minnesota Press, Copyright 1962 de la Universidad de Minnesota. Extracto impreso con el permiso del editor.

"Explanation in Science and in History", en R. Colodny (comps.): *Frontiers of Science and Philosophy*. Pittsburgh, University of Pittsburgh Press, 1962, págs. 9-33. Extracto impreso con el permiso del editor.

"Rational Action", en: *Proceedings and Addresses of the American Philosophical Association*, vol. 35, 1961-1962, págs. 5-23. Yellow Springs, Ohio, The Antioch Press, 1962. Extracto impreso con el permiso de American Philosophical Association.

¿Cuál es la naturaleza de las explicaciones que puede brindar la ciencia empírica? ¿Qué comprensión de los fenómenos empíricos transmiten? Este ensayo trata de aclarar estas cuestiones examinando en detalle la forma y la función de algunos de los principales tipos de explicaciones que han sido ofrecidas en diferentes campos de la ciencia empírica.

Entenderemos aquí las expresiones "ciencia empírica" y "explicación científica" en un sentido que abarca todo el ámbito de la investigación empírica, incluidas tanto las ciencias naturales y sociales como la investigación histórica. Este uso amplio de las dos expresiones no pretende prejuzgar sobre la cuestión referida a las semejanzas y diferencias lógicas y metodológicas entre campos diferentes de la investigación empírica, sino indicar que los procedimientos utilizados en esos campos diferentes deben ajustarse a ciertas normas básicas de objetividad. De acuerdo con estas normas, las hipótesis y las teorías —incluyendo aquellas que se invocan con propósitos explicativos— deben ser testables con referencia a elementos de juicio públicamente discernibles, y su aceptación estará siempre sujeta a la condición de que puedan abandonarse si se encuentran elementos de juicio adversos o hipótesis o teorías más adecuadas.

Podemos considerar una explicación científica como una respuesta a una pregunta concerniente al porqué, por ejemplo: "¿Por qué los planetas se mueven en órbitas elípticas, uno de cuyos focos lo ocupa el sol?", "¿Por qué la luna parece mucho mayor cuando está cerca del horizonte que cuando está alta en el cielo?", "¿Por qué falló el aparato de televisión del Ranger VI?", "¿Por qué los hijos de padres con ojos azules tienen siempre ojos azules?", "¿Por qué Hitler declaró la guerra a Rusia?" Hay otras maneras de formular lo que llamaremos *preguntas que piden una explicación*: podemos preguntar qué fue lo que causó el fracaso del aparato de televisión del Ranger IV o qué fue lo que condujo a Hitler a su funesta decisión. Pero una pregunta de la forma "¿Por qué...?" siempre constituye una formulación típica adecuada, aunque a veces parezca un poco torpe.

A veces se indica el objeto de una explicación, o *explanandum*, mediante un sustantivo, por ejemplo, cuando pedimos una explicación de la aurora boreal. Es importante comprender que este tipo de formulación sólo tiene un significado claro siempre que se la pueda reformular en términos de una pregunta "¿por qué...?" Así, en el contexto de una explicación debe considerarse que la aurora boreal se caracteriza por ciertos rasgos generales distintivos, cada uno de los cuales puede describirse mediante una cláusula "que...", por ejemplo: que sólo se la encuentra normalmente en latitudes septentrionales bastante altas; que aparece intermitentemente; que los máximos de las manchas solares, con su ciclo de once años, están acompañados regularmente por máximos en la frecuencia y el brillo de la aurora boreal; que una aurora presenta líneas espectrales características correspondientes a gases atmosféricos raros, etc. Y pedir una explicación de la aurora boreal equivale a solicitar una explicación de *por qué* sus manifestaciones se producen de la manera indicada y *por qué* éstas tienen características físicas como las que acabamos de mencionar. En verdad, pedir una explicación de la aurora boreal, de las mareas,

de los eclipses solares en general o de un eclipse solar determinado, o de una determinada epidemia de gripe, etc., sólo tiene un significado claro si se entiende cuáles aspectos del fenómeno en cuestión deben explicarse; y, en este caso, el problema que se quiere explicar puede también expresarse en la forma "¿por qué se da el caso de que *p*?", donde el lugar de "*p*" está ocupado por un enunciado empírico que especifica el explanandum. Las preguntas de este tipo serán llamadas *preguntas sobre el porqué que piden una explicación.*

Sin embargo, no todas las preguntas sobre el porqué piden explicaciones. Algunas de ellas solicitan razones en apoyo de una aserción. Así, enunciados tales como "el huracán Dalila virará hacia el Atlántico", "debe de haber muerto de un ataque al corazón", "a Platón no le habría gustado la música de Stravinsky" podrían provocar la pregunta "¿por qué debe ser así?", que pide no una explicación, sino pruebas, fundamentos o razones en apoyo de la aserción expresada. Las preguntas de este tipo serán llamadas *preguntas que piden razones* o *preguntas epistémicas.* Expresarlas en la forma "¿por qué debe darse el caso de que *p*?" es engañoso; es más adecuado transmitir su intención mediante una formulación tal como "¿por qué debe creerse que *p*?" o "¿qué razones hay para creer que *p*?"

Una pregunta sobre el porqué que pide una explicación normalmente presupone que el enunciado que ocupa el lugar de "*p*" es verdadero, y solicita una explicación del hecho, suceso o estado de cosas presunto descripto por él. Una pregunta epistémica sobre el porqué no presupone la verdad del enunciado correspondiente, sino que, por el contrario, solicita razones para creer en su verdad. Una respuesta adecuada a la primera, pues consistirá en ofrecer una explicación de un presunto fenómeno empírico, mientras que una respuesta adecuada a la segunda ofrecerá razones que convaliden o justifiquen un enunciado. A pesar de estas diferencias en las presuposiciones y los objetivos, hay también importantes conexiones entre los dos tipos de preguntas; en particular, como expondremos más adelante (secciones 2.4 y 3.5), toda respuesta adecuada a una pregunta que pide una explicación, "¿por qué se da el caso de que *p*?", también suministra una respuesta potencial a la correspondiente pregunta epistémica "¿qué fundamentos hay para creer que *p*?"

En el examen siguiente, estableceré primero una distinción entre dos tipos básicos de explicación científica, el deductivo-nomológico y el inductivo-estadístico, cada uno de los cuales está caracterizado por un "modelo" esquemático; y examinaré ciertas cuestiones lógicas y metodológicas que plantean estos modelos, incluyendo una serie de objeciones que han sido dirigidas contra ellos. Después de esto, me propongo evaluar la significación y la corrección de las concepciones básicas inherentes a esos modelos, explorando en qué medida pueden éstos servir para analizar la estructura y aclarar el fundamento teórico de diferentes tipos de explicaciones que encontramos en las ciencias empíricas.

2. *La explicación deductivo-nomológica*

2.1. *Cuestiones básicas: la explicación D-N y el concepto de ley*

En su libro *How We Think*,[2] John Dewey describe un fenómeno que observó un día mientras estaba lavando platos. Después de sacar unos vasos de vidrio del agua jabonosa caliente y de colocarlos boca abajo sobre una plancha, observó que aparecían pompas de jabón debajo de los bordes de los vasos, crecían por un momento, luego se detenían y finalmente volvían a entrar en los vasos. ¿Por qué sucedía esto? Dewey esbozó la siguiente explicación: al trasladar los vasos a la plancha, entraba aire frío en ellos; el aire era calentado de modo gradual por el vidrio, que tenía inicialmente la temperatura del agua caliente jabonosa. Esto daba origen a un aumento del volumen del aire contenido en el vaso y, de este modo, a la dilatación de la película de jabón formada entre la plancha y los bordes de los vasos. Pero luego el vidrio se enfriaba gradualmente, al igual que el aire contenido en su interior, como resultado de lo cual las pompas de jabón retrocedían.

La explicación esbozada puede considerarse como un razonamiento según el cual el fenómeno que se quiere explicar, *el fenómeno del explanandum,* es de esperar en virtud de ciertos hechos explicativos. Estos pueden clasificarse en dos grupos: (a) hechos particulares, y (b) uniformidades expresables por medio de leyes generales. El primer grupo incluye hechos como los siguientes: los vasos fueron sumergidos en agua jabonosa de temperatura considerablemente superior a la del aire circundante; fueron colocados boca abajo sobre una plancha, sobre la cual se formó un charquito de agua jabonosa que dio origen a la película de jabón, etc. El segundo grupo de hechos explicativos estaría constituido por las leyes sobre los gases y varias otras concernientes al intercambio de calor entre cuerpos de diferente temperatura, a la conducta elástica de las pompas de jabón, etc. Mientras que algunas de estas leyes sólo son insinuadas por formulaciones tales como "el calentamiento del aire del vaso provocó el aumento de su presión", y otras ni siquiera son mencionadas de esta manera indirecta, se las presupone claramente en la afirmación de que ciertas etapas del proceso dan origen a otras. Si imaginamos expresadas por completo las diversas suposiciones explicativas tácitas o implícitas, entonces la explicación puede concebirse como un razonamiento deductivo de la forma:

$$\text{(D-N)} \qquad \frac{\left.\begin{array}{l} C_1, C_2, \ldots, C_k \\ L_1, L_2, \ldots, L_r \end{array}\right\} \text{Explanans } S}{E \qquad \text{Oración-explanandum}}$$

Aquí, C_1, C_2, ..., C_k son oraciones que describen los hechos particulares aducidos; L_1, L_2, ..., L_r son las leyes generales sobre las que reposa la explicación. Se dirá que estas oraciones, en conjunto, constituyen el *explanans S,* donde *S* puede ser concebido alternativamente como el con-

[2] Dewey (1910), cap. VI.

junto de las oraciones explicativas o como su conjunción. La conclusión E del razonamiento es una oración que describe el fenómeno aludido en el explanandum; la llamaré oración-explanandum o enunciado-explanandum. La palabra "explanandum" sola se utilizará para significar el fenómeno del explanandum o la oración-explanandum: el contexto indicará en cada caso cuál de esos significados tiene.

El tipo de explicación cuya estructura lógica está indicada por el esquema (D-N) se denominará explicación *deductivo-nomológica*, o explicación D-N, para abreviar, pues realiza una subsunción del explanandum bajo principios que tienen el carácter de leyes generales. Así, una explicación D-N responde a la pregunta "*¿por qué* se produjo el fenómeno mencionado en el explanandum?", mostrando que dicho fenómeno resultó de ciertas circunstancias particulares, especificadas en C_1, C_2, ..., C_k, de acuerdo con las leyes L_1, L_2, . . ., L_r. Al señalar esto, el razonamiento demuestra que, dadas las circunstancias particulares y las leyes en cuestión, la aparición del fenómeno *era de esperar*; y es éste el sentido en el cual la explicación nos permite *comprender por qué* se produjo el fenómeno.[3]

En una explicación D-N, pues, el explanandum es una consecuencia lógica del explanans. Además, en una explicación D-N las leyes generales son esenciales; es en virtud de ellas por lo que los hechos particulares citados en el explanans poseen atingencia explicativa al fenómeno del explanandum. Así, en el caso de las pompas de jabón de Dewey, el gradual calentamiento del aire frío contenido en los vasos calientes constituiría un mero antecedente accidental, más que una factor explicativo de las pompas, de no ser por las leyes sobre los gases que vinculan los dos hechos. Pero, ¿qué sucede si la oración-explanandum E, en un razonamiento de la forma (D-N), es una consecuencia lógica de las oraciones C_1, C_2, . . ., C_k solamente? En tal caso, sin duda, no se *necesitan* leyes empíricas para deducir E del explanans; y toda ley contenida en éste es una premisa gratuita y prescindible. Esto es verdad; pero en tal caso, el razonamiento no sería una explicación. Por ejemplo, el razonamiento:

3 Varios pensadores del pasado adoptaron una concepción general de la explicación científica según la cual ésta supone una subsunción deductiva en las leyes generales, aunque no siempre la formularon claramente, y también ha sido defendida por varios autores contemporáneos, entre ellos N. R. Campbell [(1920), (1921)], quien desarrolló esa idea con detalle considerable. En un libro de texto publicado en 1934, esa concepción fue formulada concisamente del siguiente modo: "La explicación científica consiste en subsumir en alguna regla o ley que exprese un carácter invariante de un grupo de sucesos, los hechos particulares cuya explicación se le atribuye. Las leyes mismas pueden explicarse de igual manera, mostrando que son consecuencias de teorías más generales" (Cohen y Nagel, 1934, pág. 397). Popper ha expuesto esta concepción de la explicación en varias de sus publicaciones; véase la nota al final de la sección 3 en Hempel y Oppenheim (1948). Su primera enunciación aparece en la sección 12 de su libro (1935), del cual su obra (1959) es una versión inglesa ampliada. Su libro (1962) contiene observaciones adicionales sobre la explicación científica. Otras referencias sobre otros en defensa de esta idea general se econtrarán en Donagan (1957), nota 2; Scriven (1959), nota 3. Sin embargo, como se verá en la sección 3, la subsunción en leyes generales no constituye la única forma de explicación científica.

Las pompas de jabón primero crecieron y luego disminuyeron
____
Las pompas de jabón primero crecieron

si bien es deductivamente válido, es evidente que no puede considerarse como una explicación de por qué las pompas primero crecieron. La misma observación se aplica a todos los otros casos de este tipo. Una explicación D-N tiene que contener en su explanans algunas leyes generales que son *necesarias* para la deducción del explanandum, es decir, cuya eliminación quitaría validez al razonamiento.

Si el explanans de una explicación D-N determinada es verdadero, vale decir, si la conjunción de sus oraciones constituyentes es verdadera, diremos que la *explicación es verdadera*; una explicación verdadera, claro está, tienen también un explanandum verdadero. Diremos también que *una explicación D-N se halla más o menos firmemente apoyada o confirmada* por un conjunto dado de elementos de juicio según que su explanans esté más o menos firmemente confirmado por dichos elementos de juicio. (Un factor que es menester considerar al juzgar la corrección empírica de una explicación determinada es la medida en la cual su explanans recibe el apoyo de los elementos de juicio totales disponibles.) Finalmente, entenderemos por una *explicación potencial D-N* todo razonamiento que tenga el carácter de una explicación D-N con la diferencia de que las oraciones que constituyen su explanans no necesitan ser verdaderas. En una explicación potencial D-N, pues, L_1, L_2, ..., L_r serán lo que Goodman ha llamado *oraciones legales*, es decir, oraciones que son similares a las leyes excepto en que pueden ser falsas. A las oraciones de este tipo también las llamaremos *nómicas o nomológicas*. Usamos la noción de explicación potencial, por ejemplo, cuando preguntamos si una ley o teoría nueva y aún no puesta a prueba brindará una explicación de un cierto fenómeno empírico; o cuando decimos que la teoría del flogisto, si bien ha sido descartada, suministró una explicación de ciertos aspectos de la combustión.[4] Hablando en términos estrictos, sólo pueden considerarse como leyes los enunciados legales verdaderos, pues no hablamos de leyes falsas de la naturaleza. Pero por conveniencia ocasionalmente usaré el término "ley" sin suponer que la oración en cuestión es verdadera, como ya he hecho en la oración precedente.

La caracterización de las leyes como oraciones legales verdaderas plantea el importante y desconcertante problema de dar una caracterización clara de las oraciones legales sin utilizar el concepto de ley. Este problema ha resultado ser sumamente difícil, y aquí sólo haré algunas observaciones sobre ciertos aspectos que son atinentes también al análisis de la explicación científica.

Las oraciones legales pueden tener muchas formas lógicas diferentes. Algunos paradigmas de oraciones nómicas, tal como "todos los gases se dilatan cuando se los calienta a presión constante", pueden concebirse como de la forma condicional universal simple "$(x)(Fx \supset Gx)$"; otros

[4] El papel explicativo de la teoría del flogisto se describe en Conant (1951), págs. 164-71. El concepto de explicación potencial fue introducido en el libro de Hempel y Oppenheim (1948), sección 7. El concepto de oración legal, en el sentido aquí indicado, se debe a Goodman (1947).

contienen una generalización tanto universal como existencial, como la oración "para todo compuesto químico, existe una gama de temperaturas y presiones a las cuales el compuesto es líquido"; muchas de las oraciones legales y de los principios teóricos de las ciencias físicas afirman relaciones matemáticas más o menos complejas entre variables cuantitativas diferentes.[5]

Pero las oraciones legales no pueden caracterizarse en función de su forma solamente. Por ejemplo, no todas las oraciones de la forma condicional universal simple que acabamos de mencionar son legales; por ende, aunque sean verdaderas, no son leyes. Las oraciones "todos los miembros de la Junta Directiva del Colegio Greenbury son calvos" y "todas las peras de esta cesta son dulces" ilustran esta afirmación. Goodman[6] ha señalado una característica que distingue a las leyes de oraciones como las anteriores: las primeras pueden dar origen a enunciados condicionales contrafácticos y subjuntivos, mientras que las segundas no. Así, la ley sobre la dilatación de los gases puede servir para dar apoyo a enunciados tales como "si se hubiera calentado (se calentara) el oxígeno contenido en este cilindro bajo presión constante, entonces se habría dilatado (se dilataría)"; mientras que el enunciado acerca de la Junta Directiva del Colegio no da ningún apoyo al condicional subjuntivo "si Robert Crocker fuera miembro de la Junta Directiva del Colegio Greenbury entonces sería calvo".

Podríamos agregar que los dos tipos de oraciones difieren, análogamente, en poder explicativo. La ley de los gases, en combinación con datos particulares adecuados, como que el oxígeno del cilindro fue calentado bajo presión constante, puede servir para explicar por qué aumentó el volumen del gas; pero el enunciado acerca de la Junta Directiva, combinado análogamente con un enunciado tal como "Harry Smith fue miembro de la Junta Directiva del Colegio Greenbury en 1964" no permite explicar por qué Harry Smith es calvo.

Pero aunque estas observaciones aclaran el concepto de legalidad, en cambio no suministran ninguna explicación satisfactoria de él. Pues una de ellas presupone la comprensión de los enunciados contrafácticos y subjuntivos que presentan notorias dificultades filosóficas, mientras que la otra utiliza la idea de explicación para aclarar el concepto de enunciado legal, y aquí estamos tratando, a la inversa, de caracterizar un cierto tipo de explicación con ayuda de conceptos entre los que se cuenta el de enunciado legal.

Ahora bien, nuestros ejemplos de oraciones no legales comparten una característica que parecería suministrar un criterio para la distinción que estamos tratando de establecer; a saber, cada uno de ellos sólo se aplica a

[5] Cuando Scriven habla "del modelo deductivo, con su forma silogística, en el que ningún estudiante de lógica elemental puede dejar de completar la inferencia, dada la premisa" (1959, pág. 462), impone al modelo una concepción muy simplificada que carece totalmente de fundamento; pues el esquema (D-N) permite claramente el uso de leyes generales muy complejas del tipo especificado en el texto; y cuando éstas aparecen en el explanans, el explanandum, por supuesto, no puede ser deducido por métodos silogísticos.

[6] Goodman (1955, pág. 25); véase ciertas reservas en *ibíd.*, pág. 118.

un número finito de casos o ejemplos individuales. ¿Y no debe considerarse que una ley general admita un número indefinido de casos?

Sin duda, una oración legal no debe estar limitada a un número finito de casos: no debe ser equivalente a una conjunción finita de oraciones singulares o, dicho brevemente, debe tener una *forma esencialmente general.* Así, la oración "todo elemento de la clase compuesta por los objetos *a*, *b* y *c* tienen la propiedad *P*" no es legal, pues es lógicamente equivalente a la conjunción "*Pa· Pb· Pc·*", y evidentemente una oración de este tipo no puede dar apoyo a condicionales contrafácticos ni suministrar explicaciones.[7]

Pero nuestras dos generalizaciones anteriores no legales no quedan excluidas por esta condición: ellas no son lógicamente equivalentes a conjunciones finitas correspondientes ya que no enuncian específicamente quiénes son los miembros de la Junta Directiva, o cuáles peras se encuentran en la cesta. ¿Debemos, pues, negar carácter legal también a toda oración general que —por accidente empírico, por así decir— sólo tiene un número finito de casos? Seguramente esto sería inadecuado. Supongamos, por ejemplo, que a partir de las leyes básicas de la mecánica celeste se infiere un enunciado general concerniente al movimiento relativo de las componentes de una estrella doble, en el caso especial en que las componentes tengan exactamente igual masa. ¿Este enunciado solo debe considerarse una ley si se ha demostrado que existen al menos dos (o quizá más) casos de este tipo especial de estrella doble? O consideremos el enunciado general, derivable de las leyes newtonianas de la gravedad y del movimiento, referente, de manera similar a la ley de Galileo, a la caída libre de cuerpos físicos cerca de la superficie, de una masa esférica que tiene la misma densidad que la Tierra, pero un radio igual al doble del de ésta. ¿Debe negarse el carácter de ley a este enunciado, a menos que se demuestre que abarca varios casos, aun cuando sea una consecuencia lógica de un conjunto de leyes con muchos casos?

Además, sólo parece haber una "diferencia de grado" no esencial entre un enunciado general que sólo tiene un ejemplo y otro que tiene dos o cualquier número finito de ejemplos. Pero, entonces, ¿cuál sería el número de ejemplos que debería tener una ley? Afirmar cualquier número finito particular sería arbitrario; y el requisito de que debe haber infinitos ejemplos reales plantearía obvias dificultades. Evidentemente, pues, el concepto de ley científica no puede someterse razonablemente a ninguna

[7] En tales referencias a "la forma" de una oración, surge otra dificultad: esa forma sólo está claramente determinada si la oración se halla expresada en un lenguaje formalizado. Una oración castellana tal como "este objeto es soluble en agua" puede concebirse como una oración singular de la forma "Pa", pero también, alternativamente, como una oración de forma general según la cual si en un momento cualquiera se coloca el objeto en una cantidad (suficientemente grande) de agua, se disolverá. (Desarrollaremos esto más adelante, en la sección 2.3.1.) Nuestra observación acerca de una oración de la forma "para todo *x*, si *x* es *a*, *b* o *c*, entonces *x* tiene la propiedad *P*" podría enunciarse con mayor circunspección diciendo que este tipo de oración no es una ley formulada *en términos de P*; no puede servir para eplicar la aparición de *P* en ningún caso particular, ni puede brindar apoyo a condicionales contrafácticos o subjuntivos acerca de apariciones particulares de *P*.

condición concerniente al número de casos, con excepción del requisito que excluye la equivalencia lógica con enunciados singulares.

Por otra parte, debemos observar que el concepto, propuesto en el examen anterior, de "caso" o "ejemplo" de un enunciado general no es en modo alguno tan claro como parece. Consideremos, por ejemplo, enunciados generales de la forma "todos los objetos que tienen la propiedad *F* tienen también la propiedad *G*" o, brevemente, "todos los *F* son *G*". Parece natural admitir el criterio de que un objeto particular *i* sea un caso de tal enunciado si, y sólo si, *i* tiene la propiedad *F* y la propiedad *G* o, brevemente, si *i* es al mismo tiempo *F* y *G*. Esto implicaría que si no hay objetos que poseen las propiedad *F*, el enunciado general no tendría ejemplos. Sin embargo, el enunciado es lógicamente equivalente a "todos los no-*G* son no-*F*", el cual, según el criterio mencionado, bien puede tener ejemplos aunque no haya ningún *F*. Así, el enunciado general "todos los unicornios se alimentan de tréboles" no tendría ejemplos, pero su equivalente "todo lo que no se alimenta de tréboles no es un unicornio" tendría muchos ejemplos, quizá infinitos. Una observación análoga podría ser cierta de la ley mencionada antes concerniente a las estrellas dobles cuyas componentes tienen igual masa. Por consiguiente, el criterio de ilustración aducido, que parece obvio a primera vista, tiene la consecuencia de que de dos enunciados generales lógicamente equivalentes, uno podría no tener ejemplos y el otro tenerlos infinitos. Esto hace inaceptable dicho criterio, pues tales oraciones equivalentes expresan la misma ley y, por ende, deben estar ilustradas por los mismos objetos.

Con respecto a las leyes del tipo simple que estamos considerando, la siguiente definición alternativa de ilustración bastará para asignar los mismos casos a enunciados equivalentes: un objeto *i* es un ejemplo del enunciado "todos los *F* son *G*" si, y sólo si, no se da el caso de que *i* sea *F* pero no *G*. Sin embargo, con respecto a leyes de forma lógica más compleja, el concepto de ejemplo plantea problemas adicionales.[8] Pero no hace falta que los investiguemos aquí, pues no sostengo que una ley deba satisfacer ciertas condiciones mínimas concernientes al número de casos.

Hay ambién otra característica común de nuestras generalizaciones no legales que parece poder suministrar un criterio para la distinción que estamos examinando: todas ellas contienen términos tales como "esta cesta" y "la Junta Directiva del Colegio Greenbury", que directa o indirectamente se refieren a objetos, personas o lugares particulares; en cambio, los términos que aparecen en las leyes de Newton o en las leyes sobre los gases no contienen tales referencias. En un artículo anterior sobre este tema,

[8] Estas dificultades concernientes a la idea intuitiva de ilustración de una ley general están estrechamente relacionadas con las paradojas de la confirmación expuestas en Hempel (1945). La siguiente consecuencia ilustra también lo inadecuado del criterio intuitivo considerado inicialmente: la oración "todos los *F* son *G*" es lógicamente equivalente a "todo lo que sea *F* pero no es *G* es *G* y no *G*"; y según el criterio aludido, esta oración evidentemente no puede tener ningún ejemplo, aunque "todos los *G* son *G*" sea verdadera y esté ilustrada por una infinita cantidad de objetos que sean *F* y *G*. Nuestro criterio modificado de ilustración evita esta dificultad: los conjuntos de ejemplos así concebidos, de dos oraciones cualesquiera con cuantificadores universales, lógicamente equivalentes y de una variable son idénticos.

Oppenheim y yo sugerimos, por ello, que los predicados constituyentes de las que llamábamos oraciones legales fundamentales deben ser todos tales que la especificación de su significado no exija referencia alguna a ningún objeto o locación particular.[9] Observábamos, sin embargo, que esta caracterización aún no es satisfactoria para propósitos de explicación porque la idea de "el significado" de un término está lejos de ser clara.

Además, la referencia a individuos particulares no siempre priva a un enunciado general de poder explicativo, como lo muestra la ley de Galileo sobre la caída libre, cuya formulación completa hace referencia a la Tierra. Ahora bien, es cierto que, con reservas que pronto enunciaremos, la ley de Galileo puede considerarse como derivable de las leyes de la teoría newtoniana que tienen el carácter de oraciones legales fundamentales, de modo que una explicación basada en la ley de Galileo también puede ser dada por medio de leyes fundamentales. Pero es indudable que no puede darse por supuesto que todas las otras leyes que mencionan individuos particulares puedan derivarse análogamente de leyes fundamentales.

Goodman, en un penetrante examen del concepto de ley, ha sostenido que a diferencia de las generalizaciones no legales las oraciones legales pueden recibir apoyo de casos observados y, por ende, ser "proyectadas" desde los casos examinados a los no examinados. Ha sostenido, además, que la "proyectabilidad" relativa de las generalizaciones está determinada principalmente por el "reforzamiento" [*entrenchment*] relativo de sus predicados constituyentes, es decir, por la medida en la cual esos predicados han sido usados en generalizaciones proyectadas previamente.[10] De este modo, los términos como "miembro de la Junta Directiva del Colegio Greenbury en 1964" y "pera de esta cesta" quedarían descalificados para los propósitos de formular oraciones legales, sobre la base de que carecen de un adecuado reforzamiento.

Pero si bien el criterio de Goodman logra excluir de la clase de las oraciones legales generalizaciones como las de nuestros dos ejemplos, la clase de las oraciones legales que delimita parece demasiado amplia para nuestros propósitos. Pues, según Goodman, el "reforzamiento de un predicado resulta de la proyección real, no solamente de este predicado, sino también de todos los predicados de la misma extensión. En cierto sentido, lo que se refuerza no es la palabra misma sino la clase que determina..."[11] Por consiguiente, reemplazar un predicado en una oración legal por otro de igual extensión daría nuevamente una oración legal. ¿Sucede esto en general? Supongamos que la hipótesis h: "$(x)(Px \supset Qx)$" tiene forma legal, pero que, de hecho, hay exactamente tres elementos en la clase deter-

[9] Hempel y Oppenheim (1948), sección 6. La "especificación de significado" podría concebirse como realizada mediante una definición o quizá por medios más débiles, tales como las oraciones de reducción de Carnap. Véase Carnap (1938) y para mayores detalles (1936-1937). La distinción así esbozada entre aquellos términos que de alguna manera se refieren a individuos y los que no tienen tal referencia es muy afín a la distinción hecho por Popper, en la sección 14 de 1935 y 1959, entre conceptos individuales, "en cuya definición son indispensables los nombres propios (o signos equivalentes)", y conceptos universales, para los que no se da tal situación.

[10] Para mayores detalles y para consideraciones adicionales acerca de la proyectibilidad, véase Goodman (1955), en particular los capítulos III y IV.

[11] Goodman, 1955, págs. 95-96.

minada por "P", a saber, a, b y c. Luego, "Px" tiene la misma extensión que "$x = a \vee x = b \vee x = c$". El reemplazo de "$Px$" por esta expresión, sin embargo, convierte a h en la oración "$(x)[(x = a \vee x = b \vee x = c) \supset Qx]$", la cual, por ser lógicamente equivalente a "$Qa \cdot Qb \cdot Qc$" no es legal, según nuestra idea de que una oración legal debe ser de forma esencialmente general para que pueda desempeñar una función explicativa. Nuestra concepción de la legalidad difiere en este punto de la de Goodman, quien introduce la noción principalmente en un esfuerzo por trazar una línea divisoria entre las oraciones confirmables por sus ejemplos y las que no lo son.[12] Quizá no sea necesario exigir de las primeras que sean de forma esencialmente general, y Goodman no impone este requisito a las oraciones legales. Sin embargo, con respecto a las leyes que deben cumplir una función explicativa, el requisito me parece indispensable.

Aunque el examen anterior no ha conducido a una caracterización general totalmente satisfactoria de las oraciones legales y, por ende, de las leyes, espero que habrá aclarado en cierta medida el sentido en el cual entenderemos esos conceptos en el presente estudio.[13]

Los ejemplos que hemos considerado hasta ahora ilustran la explicación deductiva de hechos particulares por medio de leyes empíricas. Pero la ciencia empírica plantea la pregunta "¿por qué?" también con respecto a las uniformidades expresadas por tales leyes, y a menudo las responde, asimismo, por medio de una explicación deductivo-nomológica en la cual subsume la uniformidad en cuestión bajo leyes más inclusivas o bajo principios teóricos. Por ejemplo, las preguntas de por qué los cuerpos en caída libre se mueven de acuerdo con la ley de Galileo, y por qué el movimiento de los planetas presenta las uniformidades expresadas por las leyes de Kepler se responden mostrando que estas leyes sólo son consecuencias especiales de las leyes newtonianas de la gravedad y del movimiento. Análogamente, las uniformidades expresadas por las leyes de la óptica geométrica, como las de la propagación rectilínea de la luz y de la reflexión y la refracción, se explican subsumiéndolas bajo los principios de la óptica ondulatoria. Para mayor brevedad, a veces nos referimos elípticamente a una explicación de una uniformidad expresada por una ley como a una explicación de la ley en cuestión.

Debe observarse, sin embargo, que en los ejemplos mencionados, la teoría invocada no implica, hablando estrictamente, las presuntas leyes generales que se quiere explicar; más bien implica que esas leyes sólo son válidas dentro de un ámbito limitado, y aun dentro de éste sólo aproximadamente. Así, la ley newtoniana de la gravedad implica que la aceleración de un cuerpo en caída libre no es constante, como afirma la ley de Galileo, sino que sufre un aumento muy pequeño pero constante a medida que el cuerpo se acerca al suelo. Pero si bien, hablando estrictamente, la ley de Newton contradice la de Galileo, también muestra que ésta se cumple casi exactamente en la caída libre en distancias cortas. Con un poco más

[12] Acerca de esta distinción, véase el agregado al artículo "Estudios en la lógica de la confirmación."

[13] Otros análisis de los problemas aquí mencionados se encontrarán en Braithwaite (1953), cap. IX y Nagel (1961), cap. 4.

de detalle, podríamos decir que la teoría newtoniana de la gravedad y el movimiento implica sus propias leyes concernientes a la caída libre en circunstancias diversas. Según una de ellas, la aceleración de un objeto pequeño que cae libremente hacia un cuerpo esférico homogéneo varía en proporción al cuadrado de su distancia del centro de la esfera y, por consiguiente, aumenta en el curso de la caída; y la uniformidad expresada por esta ley queda explicada, en un sentido estrictamente deductivo, por la teoría newtoniana. Pero cuando se la une a la afirmación de que la Tierra es una esfera homogénea de masa y radio específicos, la ley en cuestión implica que para la caída libre en distancias cortas cercanas a la superficie de la Tierra, la ley de Galileo se cumple con un elevado gasto de aproximación; en este sentido, podría decirse que la teoría suministra una *explicación D-N aproximada* de la ley de Galileo.

De igual modo, en el caso del movimiento planetario, la teoría newtoniana implica que puesto que un planeta está sujeto a la atracción gravitacional tanto del sol como de los otros planetas, su órbita no será exactamente elíptica, sino que presentará ciertas perturbaciones. Por lo tanto, como observó Duhem,[14] la ley newtoniana de la gravedad, lejo de ser una generalización inductiva basada en las leyes de Kepler, es incompatible con ellas, hablando estrictamente. Uno de sus méritos importantes es, precisamente, el hecho de que permite al astrónomo calcular las desviaciones de los planetas de las órbitas elípticas que Kepler les había asignado.

Una relación similar existe entre los principios de la óptica ondulatoria y las leyes de la óptica geométrica. Por ejemplo, la primera exige una "curvatura" de difracción de la luz alrededor de los obstáculos, fenómeno excluido por la concepción de la luz como compuesta de rayos que se propagan en línea recta. Pero, en analogía con el ejemplo precedente, la explicación de la teoría ondulatoria implica que las leyes de la propagación rectilínea, de la reflexión y de la refracción, tales como las formula la óptica geométrica, se cumplen con un alto grado de aproximación dentro de un ámbito limitado de casos, incluyendo aquellos que dan apoyo experimental a las leyes en su formulación original.

En general, una explicación basada en principios teóricos ampliará y profundizará nuestra comprensión de los fenómenos empíricos a los que ella se refiere. Ampliará la comprensión porque la teoría, por lo general, abarcará un ámbito mayor de sucesos que las leyes empíricas establecidas previamente. Por ejemplo, la teoría newtoniana de la gravitación y el movimiento abarca la caída libre, no solamente en la Tierra, sino también en otros cuerpos celestes; y no solamente los movimientos planetarios, sino también el movimiento relativo de las estrellas dobles, las órbitas de los cometas y de los satélites artificiales, los movimientos de los péndulos, ciertos aspectos de las mareas y muchos otros fenómenos. Y una explicación teórica profundiza nuestra comprensión al menos por dos razones.

[14] Véase Duhem, 1906, págs. 312 y sigs. Las observaciones de Duhem sobre esta cuestión están incluidas en los extractos de la traducción de P. P. Wiener de la obra de Duhem reimpresos en Feigl y Brodbeck (1953). Este punto ha sido destacado por varios autores, entre ellos Popper (1957 a, págs. 29-34) y Feyerabend (1962, págs. 46-48).

En primer término, presenta las diferentes regularidades que presentan diversos fenómenos, por ejemplo, las que acabamos de mencionar con referencia a la teoría de Newton, como manifestaciones de unas pocas leyes básicas. En segundo término, como ya hemos observado, las generalizaciones aceptadas antes como enunciaciones correctas de regularidades empíricas aparecerán, habitualmente, como aproximaciones sólo de ciertos enunciados legales implicados por la teoría explicativa, y que tiene validez dentro de cierto ámbito limitado. Y en la medida en que los tests de las leyes, en su anterior formulación, se limitaban a los casos que caen dentro de este ámbito, la explicación teórica también indica por qué esas leyes, aunque no son verdaderas en general, han hallado confirmación.

Cuando una teoría científica es superada por otra, en el sentido en el cual la mecánica y la electrodinámica clásicas fueron superadas por la teoría especial de la relatividad, por lo general la nueva teoría tiene un ámbito explicativo más amplio que incluye fenómenos no explicados por la teoría anterior; y, por lo común, brinda explicaciones aproximadas de las leyes empíricas implicadas por su predecesora. Así, la teoría especial de la relatividad implica que las leyes de la teoría clásica se cumplen con mucha aproximación en los casos en los que los movimientos implicados sólo tienen velocidades pequeñas en comparación con la de la luz.

La concepción general de la explicación por la subsunción deductiva bajo leyes generales o principios teóricos, tal como ha sido esbozada en esta sección, será llamada el *modelo deductivo-nomológico*, o el *modelo D-N de explicación*; las leyes invocadas en tal explicación también serán llamadas, adoptando la sugestiva expresión de William Dray, *leyes inclusivas*.[15] A diferencia de Dray, sin embargo, no llamaré al modelo D-N el modelo de ley abarcante, pues luego introduciré un segundo modelo básico de explicación científica que también se basa en las leyes inclusivas, pero que no es de la forma deductivo-nomológica. La expresión "modelo de ley abarcante" servirá, pues, para aludir a cualquiera de esos modelos.

Como indica claramente el esquema (D-N), no suponemos que una explicación deductivo-nomológica sólo invoca una ley abarcante; nuestros ejemplos revelan, en verdad, que en la explicación de un fenómeno pueden invocarse muchas leyes diferentes. Pero debemos hacer aquí una observación puramente lógica. Si una explicación tiene la forma (D-N), entonces las leyes L_1, L_2, ..., L_r invocadas en el explanans implican lógicamente una ley L^* que basta por sí misma para explicar el hecho mencionado en el explanandum con referencia a las condiciones particulares indicadas en las oraciones C_1, C_2, ..., C_k. Esta ley L^* establece que cuando se cumplen las condiciones del tipo descripto en las oraciones C_1, C_2, ..., C_r, entonces se produce un hecho del tipo descripto por la oración-explanandum.[16] Tomemos un ejemplo: un trozo de hielo flota en un vaso grande de agua, a la temperatura ambiente. Puesto que el hielo se extiende por encima de la superficie, cabría esperar que el nivel del agua se elevara a medida que el hielo se derritiese; en realidad, permanece igual. Este hecho puede

[15] Para el uso de Dray de los términos "ley inclusiva" y "modelo de ley inclusiva", véase Dray (1957 y también 1963, pág. 106).

[16] Esto fue observado ya en Hempel (1942, sección 2.1).

explicarse brevemente de la siguiente manera. Según el principio de Arquímedes, un cuerpo sólido que flota en un líquido desplaza un volumen de líquido que tiene el mismo peso que el cuerpo mismo. Por consiguiente, el trozo de hielo tiene tiene el mismo peso que el agua desplazada por su parte sumergida. Puesto que al derretirse el peso no cambia, el hielo se convierte en una masa de agua del mismo peso y, por ende, también del mismo volumen que el agua desplazada inicialmente por su parte sumergida. Por lo tanto, el nivel del agua no cambia. Las leyes sobre las cuales se basa esta explicación incluyen el principio de Arquímides, una ley concerniente al derretimiento del hielo a la temperatura ambiente, el principio de la conservación de la masa, etc. Ninguna de estas leyes menciona el vaso de agua particular o el trozo de hielo particular al que se refiere la explicación. Luego, las leyes no sólo implican que a medida que se derrite ese trozo particular de hielo en ese vaso particular, el nivel del agua no cambia, sino más bien el enunciado general L^* según el cual, en el mismo *tipo* de circunstancias, es decir, cuando cualquier trozo de hielo flota en el agua de cualquier vaso a la temperatura ambiente, se produce el mismo *tipo* de fenómeno, o sea, el nivel del agua no cambia. La ley L^* por lo común será "más débil" que las leyes L_1, L_2, ..., L_k; es decir, es lógicamente implicada por la conjunción de estas leyes, pero, en general, no implicará tal conjunción. Así, en nuestro ejemplo, una de las leyes explicativas originales también se aplica a la flotación de un trozo de mármol en el mercurio o de un bote en el agua, mientras que L^* sólo se refiere al caso del hielo que flota en el agua. Pero evidentemente, L^* en conjunción con C_1, C_2, ..., C_k implica por lógica E y puede utilizarse para explicar, en este contexto, el hecho descripto por E. Por lo tanto, podríamos llamar a L^* una *ley abarcante mínima* implícita en una explicación D-N dada.[17] Pero si bien tales leyes podrían usarse con propósitos explicativos, el modelo D-N no restringe en modo alguno las explicaciones deductivo-nomológicas al uso de leyes mínimas. En verdad, tal restricción no haría justicia a un objetivo importante de la investigación científica, a saber, el de establecer leyes y teorías de ámbito amplio, en el cual podrían subsumirse generalizaciones más estrechas como casos especiales o como cercanas aproximaciones a ellas.[18]

[17] No es necesario que nos detengamos en el problema de formular una definición precisa de esta noción: sólo se lo puede resolver con referencia a algún lenguaje formalizado, y para nuestros fines basta la aproximada caracterización que hemos dado. Dicho sea de paso, la noción de "el número de leyes" invocada en una explicación dada no es tan clara como podría parecer, pues a veces una ley puede formularse, de manera muy plausible, como una conjunción de dos o más leyes, y recíprocamente, a veces varias leyes pueden unirse en una. Pero repetimos que no necesitamos entrar en este problema.

[18] En uno de sus ensayos Feyerabend criticó el modelo deductivo de explicación por conducir "a la exigencia... de que todas las teorías exitosas en una dominio determinado sean consistentes entre sí" (1962, pág. 30), o con más detalle, de que sólo son admisibles (para la explicación y la predicción) en un dominio dado aquellas teorías que *contienen* a las ya usadas en este dominio o, al menos, son *consistentes* con ellas" (1962, pág. 44, bastardillas del autor). Feyerabend arguye correctamente que esta exigencia entra en conflicto con el procedimiento científico real y es inaceptable sobre fundamentos metodológicos. Pero se equivoca totalmente en su afirmación, para la que no brinda ninguna justificación, de que concebir la

2.2. *La explicación causal y el modelo D-N*

A menudo se concibe la explicación de un suceso particular como indicando aquello que lo "causó". Así, podría decirse que la dilatación inicial de las pompas de jabón descripta por John Dewey fue causada por el calentamiento del aire contenido en los vasos. Pero las atribuciones causales de esta especie presuponen leyes adecuadas, como la de que a presión constante el volumen de un gas aumenta a medida que se eleva su temperatura. Y en virtud de presuponer de este modo leyes generales que vinculan la "causa" con el "efecto", la explicación causal se ajusta al modelo D-N. Permítaseme ampliar e ilustrar brevemente esta observación.

Consideremos primero el uso explicativo de los que podría llamarse enunciados *generales* de conexión causal: éstos afirman que un hecho de cierto *tipo A* (por ejemplo, el movimiento de un imán a lo largo de una espiral de alambre cerrada) causa un hecho de otro *tipo B* (por ejemplo, el paso de una corriente eléctrica por el alambre). Sin entrar en un análisis más detallado podemos decir que en el caso más simple un enunciado de este tipo afirma una ley según la cual, cuando se produce un hecho del tipo *A*, entonces se produce un hecho correspondiente del tipo *B*, en el mismo lugar o en un lugar diferente especificable. Esto es válido, por ejemplo, para los enunciados según los cuales el movimiento de un imán causa el paso de una corriente por una espiral de alambre vecina y la elevación de la temperatura de un gas, bajo presión constante, hace aumentar su volumen. Pero muchos enunciados generales de conexión causal exigen un análisis más complejo. Así, el enunciado según el cual, en un mamífero, la detención del corazón causa la muerte, presupone ciertas condiciones "normales" no enunciadas explícitamente, pero que seguramente excluyen, por ejemplo, el uso de una máquina corazón-pulmón. "Decir que *X* causa *Y* equivale a decir que, en condiciones adecuadas, un *X*

explicación como una subsunción deductiva en leyes generales o principios teóricos implica la máxima metodológica repudiada. En realidad, el modelo D-N de explicación sólo concierne a la relación entre el explanans y el explanandum y no implica absolutamente nada acerca de la compatibilidad de los diferentes principios explicativos que podrían aceptarse sucesivamente en un campo determinado de las ciencias empíricas. En particular, no implica que una nueva teoría explicativa sólo pueda aceptarse con la condición de que sea lógicamente compatible con las aceptadas antes. Un mismo fenómeno, o conjunto de fenómenos, puede ser subsumible deductivamente en leyes y teorías distintas y lógicamente incompatibles. Ilustremos esto de manera esquemática: el hecho de que tres objetos, *a*, *b* y *c*, cada uno de los cuales tiene la propiedad *P*, posean también la propiedad *Q* puede explicarse deductivamente por la hipótesis H_1, según la cual solamente los *P* son *Q*, y alternativamente por la hipótesis H_2, según la cual todos los *P* y también algunos *no-P* son *Q*; es decir, la oración-explanandum "*Qa.Qb.Qc*" puede ser deducida de "*Pa.Pb.Pc*" en conjunción con H_1 o H_2, aunque H_1 y H_2 sean lógicamente incompatibles. Así una "nueva" teoría explicativa para determinada clase de fenómenos puede explicar deductivamente tales fenómenos aunque por lógica sea incompatible con una teoría anterior que también los explique en forma deductiva. Pero las teorías rivales no pueden ser ambas verdaderas, y bien podría ser que la anterior teoría fuera falsa. De ahí que sea realmente incorrecto el principio criticado por Feyerabend. Pero esta observación no afecta al modelo D-N de explicación, que no implica para nada tal principio.

será seguido por un *Y*", para decirlo con palabras de Scriven.[19] Cuando se usa este tipo de locución causal, habitualmente hay cierto conocimiento de cuáles son las condiciones básicas "adecuadas" o "normales" que se presuponen en el contexto dado. Pero en la medida en que esas condiciones permanecen indeterminadas, un enunciado general de conexión causal equivale, a lo sumo, a la vaga afirmación de que *hay* ciertas condiciones básicas no especificadas cuya mención explícita en el enunciado en cuestión daría una ley verdaderamente general que vinculara la "causa" con el "efecto" en cuestión.

Consideremos ahora los enunciados de conexiones causales entre sucesos individuales. Tomemos, por ejemplo, la afirmación de que la dilatación y la ulterior contracción de las pompas de jabón de Dewey fueron causadas por una elevación y una subsiguiente caída de la temperatura del aire contenido en los vasos. Evidentemente, esos cambios de temperatura sólo dan la explicación requerida en conjunción con otras condiciones, como ser la presencia de una película de jabón, una temperatura y una presión prácticamente constantes del aire exterior a los vasos, etc. Por consiguiente, en el contexto de explicación, debe admitirse que una "causa" es un conjunto más o menos complejo de circunstancias y hechos que podría describirse por un conjunto de enunciados C_1, C_2, ..., C_k. Y como lo indica el principio "a igual causa, igual efecto", la afirmación de que esas circunstancias, conjuntamente, causaron un hecho determinado implica que siempre y cuando se den las circunstancias del tipo en cuestión, se producirá un hecho del tipo que se quiere explicar. Así, la explicación causal afirma implícitamente que hay leyes generales, digamos L_1, L_2, ..., L_r, en virtud de las cuales la aparición de los antecedentes causales mencionados en C_1, C_2, ..., C_k es una condición suficiente para la aparición del hecho señalado en el explanandum. Esta relación entre factores causales y efecto se refleja en nuestro esquema (D-N): la explicación causal es, al menos implícitamente, deductivo-nomológica.

Permítaseme formular la afirmación anterior en términos más generales. Cuando se dice que un hecho individual *b* ha sido causado por otro hecho individual *a*, está implícita la afirmación de que siempre que se realiza "la misma causa" se producirá "el mismo efecto". Pero esta afirmación no debe ser entendida en el sentido de que cuando se da *a* también se da *b*; pues *a* y *b* son hechos individuales de locaciones espaciotemporales particulares, por lo cual sólo pueden darse una vez. Más bien, *a* y *b* deben considerarse como hechos particulares de ciertos *tipos* (como el calentamiento o enfriamiento de un gas, la dilatación o contracción de un gas) de los cuales puede haber muchos casos. Y la ley tácitamente implicada por la afirmación de que *b*, como hecho de tipo *B*, fue causado por *a*, como hecho de tipo *A*, es un enunciado general de conexión causal que afirma que, en condiciones adecuadas, todo caso de *A* va acompañado invariablemente por un caso de *B*. En la mayoría de las explicaciones causales no se enuncian de manera completa las circunstancias requeridas. El contenido de la afirmación de que *b* fue causado por *a* puede ser indicado, entonces, por la siguiente formulación aproximada: el suceso *b* fue

[19] Scriven, 1958, pág. 185.

precedido, de hecho, por el suceso *a* en circunstancias que, si bien no están especificadas completamente, fueron de tal tipo que la aparición de un suceso de tipo *A*, en tales circunstancias, es seguida universalmente por un suceso del tipo *B*. Por ejemplo, la afirmación de que el incendio (suceso de tipo *B*) de una parva de heno particular fue causado por un cigarrillo encendido arrojado en el heno (suceso particular de tipo *A*) asevera, ante todo, que el último suceso se produjo; pero un cigarrillo encendido incendia una parva de heno sólo si se cumplen ciertas condiciones adicionales, que por el momento no pueden enunciarse por completo; así, la atribución causal en cuestión implica, en segundo término, que se presentaron condiciones adicionales de un tipo no especificado totalmente y en las cuales un suceso de tipo *A* es seguido invariablemente por un suceso de tipo *B*.

En la medida en que un enunciado de causación individual deja indefinidas las condiciones antecedentes, y por ende también las leyes explicativas requeridas, es como una nota en la que se dice que hay un tesoro oculto en alguna parte. Su significación y su utilidad aumentarán a medida que se circunscriba con mayor precisión la ubicación del tesoro, a medida que las condiciones atinentes al caso y las leyes abarcantes correspondientes sean determinadas con creciente claridad. En algunos casos, esto se puede lograr muy satisfactoriamente; entonces, aparece la estructura de leyes abarcantes y es posible someter a prueba el enunciado que expresa la conexión causal individual. En cambio, cuando las condiciones o leyes del caso permanecen indefinidas en alto grado, la enunciación de una conexión causal tiene más bien la naturaleza de un programa, de un esbozo, de explicación en términos de leyes causales; también se la puede considerar como una "hipótesis de trabajo" que puede resultar valiosa si da una orientación nueva y fructífera a la investigación ulterior.

La concepción aquí adoptada acerca de los enunciados de causación individual puede recibir adicional claridad mediante algunos comentarios sobre la tesis de que "cuando alguien afirma que *X* causa *Y* se compromete, sin duda, con la generalización de que una causa idéntica producirá un idéntico efecto, pero en modo alguno lo obliga a aducir leyes que no contengan el término 'idéntico' para justificar tal afirmación. Aducir leyes es una manera, no necesariamente más concluyente y habitualmente menos fácil que otras, de dar apoyo al enunciado causal... (Creo que la idea de causación individual tiene esta base considerable.)" [20] Debemos distinguir claramente aquí dos cuestiones, a saber, (a) qué es lo que se afirma con el enunciado de que *X* causa *Y* (donde, en el caso de la "causación individual", *X* e *Y* son sucesos individuales), en particular, si afirmarlo compromete con una generalización; y (b) qué tipo de elementos de juicio darían apoyo al enunciado causal, en particular si sólo se lo puede sustentar aduciendo generalizaciones en forma de leyes.

En lo concerniente a la primera cuestión, he sostenido que debe interpretarse el enunciado causal dado como si afirmara por implicación que es válida una ley, o conjunto de leyes, en virtud de la cual *X* causa *Y*. Pero, como observamos antes, las leyes en cuestión *no pueden* expresarse diciendo que una causa idéntica producirá un efecto idéntico; pues si *X*

[20] Scriven, 1958, pág. 194.

e *Y* son sucesos individuales, con locaciones espaciotemporales específicas, la reaparición de una causa idéntica a *X* o de un efecto idéntico a *Y* es lógicamente imposible. La afirmación general implicada por el enunciado de causación individual según el cual *X* causó *Y* es, más bien, de ltipo sugerido en nuestro examen de la aserción de que el suceso individual *a*, como caso de *A*, causó el suceso individual *b*, como caso de *B*.

Pasemos ahora a la segunda cuestión. En ciertos casos, como el de las pompas de jabón observadas por Dewey, algunas de las leyes que vinculan los sucesos individuales *X* e *Y* pueden ser enunciables explícitamente, y entonces surge la posibilidad de aportar elementos de juicio a su favor mediante experimentos u observaciones adecuados. Por consiguiente, aunque el enunciado de conexión causal individual *afirma* implícitamente la existencia de leyes subyacentes, tal afirmación puede ser *sustentada* por elementos de juicio consistentes en casos confirmatorios particulares, en lugar de leyes generales. En otros casos, cuando la afirmación nomológica implícita en un enunciado causal indica meramente que *hay* factores atinentes a la cuestión y leyes adecuadas que conectan *X* e *Y*, puede darse cierto crédito a esta afirmación mostrando que, en ciertas condiciones, un suceso de tipo *X* está acompañado, al menos con mucha frecuencia, por un suceso de tipo *Y*; esto podría justificar la hipótesis de trabajo de que las condiciones básicas pueden ser precisadas de una manera que al final conduzcan a una conexión estrictamente causal. Es este tipo de prueba estadística, por ejemplo, el que se aduce en apoyo de afirmaciones como la de que fumar cigarrillos es "una causa de" o "un factor causante del" cáncer de pulmón. En este caso, las leyes causales presuntas no pueden formularse explícitamente por el momento. Así, la afirmación nomológica implicada por esta conjetura causal es de tipo existencial; tiene el carácter de una hipótesis de trabajo para la investigación posterior. La prueba estadística aducida presta apoyo a la hipótesis y sugiere nuevas investigaciones tendientes a determinar con mayor precisión las condiciones en las cuales fumar da origen al cáncer de pulmón.

Los mejores ejemplos de explicaciones que se ajustan al modelo D-N se basan en teorías físicas de carácter determinista. Dicho brevemente, una teoría determinista trata de los cambios de "estado" de sistemas físicos de algún tipo especificado. El estado de tal sistema en un momento dado está caracterizado por los valores que adoptan en ese momento ciertas características cuantitativas del sistema, las llamadas variables de estado. Y las leyes especificadas por tal teoría para los cambios de estado son deterministas en el sentido de que, dado el estado del sistema en un momento cualquiera, ellas determinan su estado en cualquier otro momento, anterior o posterior. Por ejemplo, la mecánica clásica brinda una teoría determinista para un sistema de masas puntuales (o cuerpos que son prácticamente pequeños en comparación con sus distancias) que se mueven sólo bajo la influencia de su mutua atracción gravitacional. Se define el estado de tal sistema en un momento dado como determinado por las posiciones y las cantidades de movimiento de sus cuerpos componentes en ese momento, sin incluir otros aspectos que puedan estar sujetos a cambio, como el color o la constitución química de los cuerpos en movimiento. La

teoría suministra un conjunto de leyes —esencialmente, las leyes newtonianas de la gravedad y del movimiento— que, dadas las posiciones y las cantidades de movimiento de los elementos de tal sistema en un momento cualquiera, determinan matemáticamente sus posiciones y cantidades de movimiento en cualquier otro instante. En particular, esas leyes permiten ofrecer una explicación D-N del hecho de que el sistema esté en un cierto estado en un momento dado, especificando en las oraciones C_1, C_2, ..., C_k del esquema (D-N) el estado del sistema en un momento anterior. La teoría mencionada ha sido aplicada, por ejemplo, a la explicación de los movimientos de los planetas y cometas, así como a los eclipses solares y lunares.

En el uso explicativo o predictivo de una teoría determinista, pues, la noción de causa como suceso antecedente más o menos bien circunscripto ha sido reemplazada por la de estado antecedente del sistema total, que suministra las "condiciones iniciales" para el cálculo, mediante la teoría, del estado posterior que se trata de explicar. Si el sistema no está aislado, es decir, si pueden actuar sobre el sistema influencias externas importantes durante el período que va del estado inicial invocado al estado que se quiere explicar, las circunstancias particulares que deben enunciarse en el explanans deben incluir esas influencias externas; y son estas "condiciones-límite", junto con las "condiciones iniciales", las que reemplazan la noción cotidiana de causa y son especificadas por los enunciados C_1, C_2, ..., C_k en la representación esquemática (D-N) de la explicación deductivo-nomológica.[21]

La explicación causal, con sus diversos grados de claridad y precisión no es, sin embargo, el único modo de explicación que se ajusta al modelo D-N. Por ejemplo, la explicación de una ley general por subsunción en principios teóricos, evidentemente, no es una explicación por causas. Pero aun cuando se las utiliza para explicar hechos individuales, no siempre las explicaciones D-N son causales. Por ejemplo, el hecho de que un péndulo simple determinado tarde dos segundos para completar una oscilación podría explicarse señalando que su longitud es de 100 centímetros y que el período t (en segundos) de cualquier péndulo simple está relacionado con su longitud l (en centímetros) por la ley: $t = 2\pi\sqrt{l/g}$, donde g es la aceleración de la gravedad. Esta ley expresa una relación matemática entre la longitud y el período (que es una característica disposicional cuantitativa) del péndulo en un mismo momento. Las leyes de este tipo, del cual son otros ejemplos las leyes de Boyle y Charles y la ley de Ohm, son llamadas a veces *leyes de coexistencia*, para distinguirlas de las *leyes de sucesión* que se refieren a los cambios temporales de un sistema. Entre estas últimas se cuentan, por ejemplo, la ley de Galileo y las leyes de los cambios de estado de sistemas incluidos en una teoría determinista. La explicación causal con referencia a hechos antecedentes presupone claramente leyes de sucesión; en el caso del péndulo, donde sólo se invoca una ley de coexistencia, segu-

[21] Para explicaciones más detalladas de las nociones de causalidad, teoría determinista y sistema determinista, véase por ejemplo, Feigl (1953); Frank (1957, capítulos 11 y 12); Margenau (1950, capítulo 19); Nagel (1961, págs. 73-78 y capítulos 7 y 10).

ramente nadie diría que el hecho de que el péndulo tenga un período de dos segundos sea *causado* por el hecho de que su longitud fuera de 100 centímetros.

Hay otro punto digno de mención. La ley del péndulo simple no sólo permite inferir el período de un péndulo a partir de su longitud, sino también, inversamente, inferir su longitud a partir de su período; en ambos casos la inferencia es de la forma (D-N). Sin embargo, una oración que exprese la longitud de un péndulo dado, junto con la ley, tendrá mucha mayor probabilidad de ser considerada una explicación del período del péndulo que la que tendría otra oración que expresara este período, junto con la ley, de ser considerada una explicación de la longitud del péndulo. Esta diferencia parece reflejar la idea de que podemos cambiar la longitud del péndulo a voluntad y, de este modo, controlar su período como "variable dependiente", mientras que el procedimiento inverso no parece posible.[22] Esta concepción es discutible, sin embargo; pues también podemos cambiar a voluntad el período de un péndulo dado, a saber, cambiando su longitud. No puede argüirse válidamente que en el primer caso el cambio de longitud sea independiente del cambio del período, pues si permanece fija la locación del péndulo, no sería posible cambiar su longitud sin cambiar también el período. En casos como éste, la concepción de sentido común de las explicaciones no parece ofrecer fundamentos claros sobre cuya base decidir si un argumento dado que subsume deductivamente un hecho bajo leyes debe ser considerado como una explicación.

En el ejemplo que acabamos de considerar, se explicaba un hecho particular, no por antecedentes causales, sino con referencia a otro hecho contemporáneo. Hasta podría argüirse que a veces puede explicarse satisfactoriamente un suceso particular con referencia a otros posteriores. Consideremos a título de ejemplo un rayo de luz que pasa de un punto *A* de un medio óptico a un punto *B* de otro medio óptico que limita con el primero a lo largo de un plano. Según el principio del tiempo mínimo de Fermat, el rayo seguirá un camino tal que la trayectoria de *A* a *B* requerirá un tiempo mínimo en comparación con cualquier otro camino. Cuál sea este camino dependerá de los índices de refracción de los dos medios, que supondremos dados. Supongamos ahora que el camino de *A* a *B* determinando por el principio de Fermat pasa a través de un punto *C* intermedio. Puede decirse entonces que este hecho tiene una explicación D-N basada en la ley de Fermat en conjunción con los datos concernientes a los medios ópticos y la información de que la luz pasó de *A* a *B*. Pero su "llegada a *B*", que sirve como uno de los hechos explicativos, sólo se produce después del hecho que se quiere explicar, a saber, el paso del rayo por *C*.

La renuencia a explicar un suceso con referencia a factores que incluyen hechos posteriores quizá surja de la idea de que las explicaciones del tipo más común, como nuestros ejemplos anteriores, parecen presentar el hecho del explanandum como habiendo sido producido por otros anteriores, mientras que no puede decirse de ningún suceso que ha sido producido por factores tales que algunos ni siquiera se habían realizado

[22] Sobre el particular, véase el examen de los enunciados causales como recetas para producir un efecto dado en Gasking (1955).

en el momento de su producción. Quizás esta idea también arroje dudas sobre las explicaciones basadas en la referencia a circunstancias simultáneas. Pero si bien tales consideraciones pueden hacer que nuestros anteriores ejemplos de explicaciones, y todas las explicaciones causales, parezcan más naturales o plausibles, no está claro qué alcance preciso puede darse a la noción de factores "que producen" un hecho dado, ni cuáles razones habría para negar el carácter de explicaciones a todas las que invoquen sucesos temporalmente posteriores al suceso que se quiere explicar.[23]

2.3. *El papel de las leyes en la explicación*

El modelo D-N, como hemos visto, asigna a las leyes o principios teóricos el papel de premisas indispensables en los razonamientos explicativos. Consideraré ahora algunas concepciones alternativas del papel de las leyes en la explicación.

2.3.1. *La concepción de las leyes como reglas de inferencia.* Una concepción que ha tenido recientemente mucha influencia considera las leyes y los principios teóricos como reglas de inferencia de acuerdo con las cuales es posible inferior enunciados particulares referentes a hechos empíricos a partir de otros enunciados semejantes.

Así, Schlick sostuvo alguna vez la opinión, que atribuía a Wittgenstein, de que "una ley natural, básicamente, no tiene el carácter lógico de una 'proposición', sino que representa 'una directiva para la formulación de proposiciones' ".[24] Schlick adhirió a esta idea en gran parte porque sostenía, entonces, que en un enunciado genuino debe poder recibir una verificación estricta de hallazgos experimentales particulares, requisito que las leyes generales evidentemente no cumplen, ya que se refiern a una cantidad indefinida de casos particulares. Pero el requisito de la verificabilidad estricta para poder considerar empíricamente significativas las proposiciones hace tiempo que ha sido abandonado por ser demasiado restrictivo,[25] y sin duda ya no constituye una buena razón para concebir las leyes como reglas y no como enunciados.

En un espíritu un poco diferente, Ryle ha caracterizado los enunciados de leyes como enunciados que pueden ser verdaderos o falsos, pero que funcionan de modo característico como licencias para la inferencia que autorizan los pasos inferenciales de la afirmación de ciertos enunciados

[23] Para otras observaciones sobre este punto, véase Scheffler (1957).

[24] Schlik (1931, pág. 190 de la traducción inglesa). Véase también la discusión de esta idea por Toulmin, quien la acepta con ciertas reservas (1953, págs. 90-105) y desarrolla, con un espíritu algo similar, una extensa analogía entre las teorías físicas y los mapas (1953, capítulo 4). Un esclarecido comentario sobre las ideas de Toulmin y sobre el problema en general se encontrará en la reseña bibliográfica hecha por Nagel del libro de Toulmin en *Mind* 63, 1954, págs. 403-412; reimpreso en Nagel, 1956, págs. 303-315.

[25] Para mayores detalles, véase el ensayo "Criterios empiristas de significación cognoscitiva; problemas y cambios", en esta serie.

fácticos a la afirmación de otros.[26] Esta concepción ha tenido influencia sobre las opiniones de otros autores acerca del papel de las leyes en la explicación científica e histórica. Dray, por ejemplo, ha hecho ciertas consideraciones interesantes en apoyo de aquélla, con especial referencia a la explicación histórica. Puesto que la explicación de un suceso histórico concreto, señala Dray, habitualmente tendrá que tomar en consideración un conjunto grande de factores atinentes a él, la ley abarcante correspondiente puede tener tantas limitaciones que sólo posea un único ejemplo, a saber, el suceso que explica. Pero en tales circunstancias, Dray plantea dudas acerca de la propiedad de aplicarle el término "ley", cuyo uso ordinario "incluye también el sentido de 'otros casos' ".[27] Sostiene, por lo tanto, que si bien al ofrecer la explicación "E porque C_1, C_2, ..., C_n", el historiador "afirma la verdad del enunciado general abarcante 'si $C_1 \dots C_n$, entonces E', ... el enunciado así obtenido ... sin duda no es más que una formulación del *principio de la inferencia del historiador,* al afirmar que puede predecirse razonablemente un resultado de este tipo a partir del conjunto de factores especificados. Puede decirse que la inferencia del historiador está *de acuerdo con* este principio. Pero es una cosa muy distinta afirmar que esta explicación implica una *ley empírica* correspondiente".[28] Para Dray, tales principios de inferencia son "condicionales generales" de la forma "si p, entonces q"; y sostiene que "pretender simplemente que un 'condicional general' está implícito en la explicación del historiador es afirmar *considerablemente menos* que lo que afirman coteóricos de las leyes abarcantes"; pues si se concibe el condicional general como una autorización para la inferencia, en el sentido de Ryle, entonces "decir que la explicación del historiador lo compromete con la 'ley' abarcante equivale simplemente a decir que lo compromete ... a razonar de manera similar en todos los otros casos *que puedan presentarse,* ya que asigna validez universal al argumento correspondiente, 'p, luego q' ".[29]

Pero, sin duda, atribuir validez universal a este esquema de razonamiento equivale a afirmar por implicación el enunciado general "siempre que p, entonces q", y viceversa: no hay ninguna diferencia en la fuerza de las afirmaciones, sino solamente en el modo de expresarlas. Y si el enunciado general tiene un solo caso, entonces lo mismo sucederá con la regla correspondiente, y se podría con igual justicia dudar de lo correcto de calificar a esta última de principio de inferencia, sobre la base de que la idea de tal principio o regla, no menos que la idea de la ley, sugiere generalidad.

En sus observaciones sobre el número de casos de una ley, Dray parece considerar que una explicación histórica es aquella que sólo utiliza un condicional general, o sea, una "ley abarcante mínima" del tipo mencionado antes. Pero, por lo común, una explicación se basará en un conjunto más o menos amplio de leyes, cada una de las cuales tiene muchos casos, y del cual la ley abarcante más restringida constituye simplemente una con-

[26] Véase Ryle (1949, págs. 121-123) y Ryle (1950).
[27] Dray, 1957, pág. 40.
[28] Dray (1957, pág. 39). Bastardillas del autor.
[29] Dray (1957, pág. 41). Bastardillas mías.

secuencia muy específica. Pero supongamos que una explicación dada reposa en sólo una generalización muy específica que tiene un único caso. ¿Puede esta generalización considerarse una ley? Nuestro examen de la sección 2.1 se refiere a esta cuestión, y bastará aquí añadir una pocas observaciones breves. Supongamos que se hace el intento de explicar la decisión de Hitler de invadir a Rusia por medio de la siguiente generalización: "Toda persona exactamente igual a Hitler en todos los aspectos, y que se enfrente exactamente con las mismas circunstancias, decide invadir a Rusia". Evidentemente, esto no suministra ninguna explicación, porque el enunciado general aducido es equivalente a la oración "Hitler decidió invadir a Rusia", que no es en absoluto una oración general y que simplemente reformula el explanandum; pues ser exactamente igual a Hitler, en todos los aspectos, es lo mismo que ser idéntico a él. Así, la generalización propuesta no tiene carácter legal porque esencialmente no está generalizada.

Pero un enunciado general, como una de las leyes abarcantes muy específicas consideradas por Dray, puede tener sólo un caso sin ser lógicamente equivalente a una proposición singular. Esta característica, como observamos antes, no priva a la generalización de su carácter legal y de su poder explicativo potencial.

Los argumentos que hemos considerado con brevedad, pues, no prestan mucho apoyo a la concepción de las leyes y los principios teóricos como reglas o principios de inferencia. Por otra parte, hay algunas consideraciones que se oponen claramente a que se los conciba de esta manera.

En primer término, en los escritos de los científicos, las leyes y los principios teóricos son tratados como enunciados. Por ejemplo, se usan enunciados generales en conjunción con enunciados singulares acerca de hechos particulares como *premisas* a partir de las cuales se infieren otros enunciados acerca de hechos particulares; análogamente, enunciados de forma general, tales como leyes de alcance más restringido, a menudo aparecen como *conclusiones* derivadas de leyes más amplias. De igual modo, las leyes generales y los principios teóricos son aceptados o rechazados sobre la base de pruebas empíricas, de manera muy semejante a los enunciados de hechos particulares, como los concernientes a la constitución del interior de la Tierra, por ejemplo.

En realidad, y esto nos lleva a una segunda dificultad, la distinción aquí presupuesta entre proposiciones singulares, por una parte, proposiciones generales, por la otra, no tiene ningún significado preciso con referencia a enunciados formulados en un lenguaje natural. Por ejemplo, el enunciado según el cual la Tierra es una esfera, puede considerarse como una proposición singular de la forma "*Et*" que asigna a un objeto particular, la Tierra, una cierta propiedad, la esfericidad. Pero también puede considerarse como un enunciado general, por ejemplo, como si afirmara que hay un punto en el interior de la Tierra con respecto al cual todos los puntos de su superficie están a la misma distancia. Análogamente, el enunciado según el cual un cristal de sal dado es soluble en agua puede considerarse como un enunciado singular que atribuye solubilidad a un objeto particular o, alternativamente, como un enunciado de carácter gene-

ral que afirma o implica que el cristal dado se disolverá en cualquier momento si se lo coloca en agua.

Puede establecerse una distinción precisa del tipo en cuestión, (*a*) si los enunciados que se quiere clasificar están expresados en un lenguaje adecuadamente formalizado que contenga una notación cuantificacional, y (*b*) si todo término no lógico del lenguaje está caracterizado como término primitivo o como término definido, y cada término definido posee una definición única basada en términos primitivos. Puede decirse que una oración de tal lenguaje es en esencia singular si es lógicamente equivalente a una oración que no contenga términos definidos ni cuantificadores; todas las otras oraciones serán esencialmente generales. La oración "la Tierra es esférica", pues, será esencialmente singular si, por ejemplo, "la Tierra" y "esférica" se consideran como términos primitivos del lenguaje en el cual están formulados nuestros enunciados; será en esencia general si, por ejemplo, "esférica" está definida por una expresión que contiene uno o más cuantificadores no eliminables.

Pero aun cuando supongamos que se ha trazado una línea divisoria precisa entre enunciados singulares y enunciados generales de esta manera o de otra similar, la propuesta de concebir los enunciados generales como reglas de inferencia para vincular enunciados singulares aún debe enfrentar otra dificultad más seria: la formulación de enunciados legales como reglas de inferencia es sumamente difícil, si no imposible, y el sistema resultante de reglas es engorroso, por decir lo menos. Sin duda, un enunciado de la forma simple "todos los F son G" o "$(x)(Fx \supset Gx)$", donde "F" y "G" son predicados primitivos en el sentido explicado, puede ser reemplazado por una regla que autorice la transición inferencial de cualquier oración de la forma "Fi" (que es singular, es decir, sin cuantificadores) a la correspondiente oración de la forma "Gi". Pero las explicaciones científicas a menudo se basan en leyes de una estructura más compleja; y para éstas, su traducción a reglas de inferencias que vinculan enunciados singulares es problemática. Por ejemplo, tomemos la ley según la cual todo metal tiene un punto de fusión específico (a la presión atmosférica), es decir, que para todo metal existe una temperatura T tal que a cualquier temperatura superior a ella el metal es sólido a la presión atmosférica. La regla de inferencia correspondiente no puede concebirse como autorizando la transición de cualquier oración de la forma "i es un metal" a la oración "hay una temperatura T tal que a cualquier temperatura inferior a ésta, pero a ninguna temperatura superior a ella, i es sólido a la presión atmosférica"; pues la conclusión así obtenida no es una oración de forma singular, sino un enunciado que contiene cuantificadores existenciales y universales. En verdad, las mismas subcláusulas "a cualquier temperatura inferior a T, i es sólido" y "a cualquier temperatura superior a T, i es no-sólido" tienen la forma universal de una ley, y la concepción general que estamos examinando parecería requerir que también ellas, a su vez, deben concebirse como reglas de inferencia y no como enunciados. Pero en el contexto dado esto no es posible, pues están calificadas por la frase cuantificadora-existencial "hay una temperatura T tal que...". En resumen, la ley dada no puede considerarse equivalente a una regla que esta-

blezca ciertas conexiones inferenciales entre oraciones singulares. Esto no significa que la ley no permita tales inferencias: en verdad, con su ayuda (es decir, utilizándola como premisa adicional) podemos inferir del enunciado "esta llave es de metal y no es líquida a los 80°C y a la presión atmosférica" otros enunciados descriptivos que afirmen que la llave no será líquida a los 74°C, los 30°C y otras temperaturas específicas inferiores a 80°C, y a presión atmosférica. Pero estas vinculaciones inferenciales y otras semejantes entre enunciados singulares realizadas por medio de la ley dada no agotan, evidentemente, su contenido, pues, como ya hemos destacado, la ley también establece conexiones, por ejemplo, entre oraciones singulares ("*i* es un metal") y oraciones cuantificadoras ("hay una temperatura *T* tal que...").

Hasta puede suceder que dos o más leyes de forma compleja, tomadas en conjunto, permitan establecer conexiones inferenciales, mientras que ninguna de ellas, por separado, permita hacerlo. Por ejemplo, dos oraciones de la forma "$(x)[Fx \supset (\exists y)Rxy]$" y "$(x)[(\exists y)(Rxy) \supset Gx]$" permiten conjuntamente la inferencia de "Gi" a partir de "Fi", pero de modo individual ninguna establece conexión alguna entre oraciones singulares. Así, la totalidad de las transiciones inferenciales entre oraciones singulares que permiten un conjunto de leyes o de principios teóricos puede exceder, en mucho, la suma (lógica o de clases) de las conexiones inferenciales establecidas entre las mismas oraciones singulares por las leyes o principios teóricos individualmente. Por consiguiente, si se insistiera en concebir las leyes y los principios teóricos científicos como reglas de inferencia extralógicas, que permitan realizar ciertas transiciones entre proposiciones singulares, entonces habría que hacerlo no para cada una de las leyes y los principios teóricos individualmente, sino para todo el conjunto de leyes y principios adoptados en un contexto dado. Sin duda, la manera más simple de hacerlo sería formular una sola regla extralógica que autorice todas aquellas transiciones —y sólo ellas— entre proposiciones singulares que pueden realizarse utilizando solamente reglas de inferencia puramente lógicas y tratando las leyes y los principios teóricos "como si" fueran proposiciones que pueden funcionar a la manera de premisas adicionales en razonamientos deductivos. Pero adoptar esta regla sería simplemente adherir sólo de manera verbal a la concepción de las leyes como reglas, y no como enunciados.[30]

[30] Es interesante observar aquí que Carnap, en su teoría de la sintaxis lógica, explícitamente admite la posibilidad de construir lenguajes con reglas extralógicas de inferencia; véase Carnap (1937, sección 51). A estas últimas las llama reglas físicas o reglas-F. Pero no afirma que todas las leyes generales o principios teóricos puedan concebirse como reglas semejantes, y subraya que la medida en la cual se deben adoptar reglas-F al construir un lenguaje es una cuestión de conveniencia. Por ejemplo, si usamos reglas-F, entonces el descubrimiento de fenómenos empíricos que entren en "conflicto" con las teorías que aceptábamos previamente nos obliga a modificar las reglas de inferencia, y, por ende, toda la estructura formal de nuestro lenguaje científico; mientras que en ausencia de reglas-F sólo es necesario efectuar una modificación de algunos enunciados teóricos aceptados con anterioridad. W. Sellars (1953) también ha propugnado la admisión de reglas materiales de inferencia en conexión con su análisis de los condicionales subjuntivos.

En suma, pues, se plantean serias dudas, por razones puramente lógicas, con respecto a si *todas* las leyes y los principios teóricos pueden ser adecuadamente concebidos como reglas de inferencia. Y aun en los casos en los que esto es posible, las consideraciones precedentes sugieren que sería más simple y más útil, para esclarecer los problemas que hemos estado examinando, concebir las leyes generales y los principios teóricos principales como enunciados; por lo tanto, adoptaremos esta posición.

2.3.2. *La concepción de las leyes como fundamento justificador de las explicaciones.* Scriven ha expuesto otra concepción que, normalmente, excluye la mención de leyes en una explicación;[31] en efecto, Scriven arguye que en la medida en que las leyes son importantes para una explicación, habitualmente funcionarán como "fundamentos justificadores" de ella. Esta concepción refleja, sin duda, la idea de que, para decirlo con palabras de Ryle, "las explicaciones no son razonamientos, sino enunciados. Son verdaderas o falsas".[32] Las explicaciones, pues, pueden tomar la forma "*q* porque *p*", donde la cláusula "*p*" menciona hechos particulares pero no leyes: y el tipo de explicación representada como un razonamiento en nuestro esquema (D-N) sería expresado por un enunciado de la forma "*E* porque C_1, C_2, ..., C_k". La citación de leyes es adecuada, según Scriven, no en respuesta a la pregunta "¿por qué *q*?", que se contesta con "*q* porque *p*", sino en respuesta a la pregunta muy diferente acerca de los fundamentos sobre los cuales los hechos mencionados en la cláusula "*p*" pueden explicar los mencionados en la cláusula "*q*". Incluir las leyes atinentes al caso en la enunciación de la explicación misma, según Scriven, sería confundir la enunciación de una explicación con la enunciación de sus fundamentos.

Ahora bien, es muy cierto que en el discurso ordinario y también en contextos científicos se responde a menudo una pregunta de la forma "¿por qué ocurrió tal hecho?" mediante un enunciado que sólo cita ciertos hechos particulares, aun en casos en los que podrían enunciarse las leyes atinentes al caso. Un ejemplo es el enunciado explicativo "el cubo de hielo se fundió porque flotaba en agua a la temperatura ambiente". Pero, como esta oración muestra también una explicación —tal como se la formula comúnmente— a menudo sólo menciona algunos de los hechos particulares de un conjunto más amplio que explican en conjunto el hecho en cuestión. No menciona otros factores que se dan por supuestos, por ejemplo, que el agua y el aire circundante permanecieron aproximadamente a la temperatura ambiente durante un tiempo adecuado. Por ende, para justificar la atribución de un papel explicativo a los hechos especificados no sólo sería necesario citar aquí ciertas leyes, sino también los hechos particulares atinentes al caso que no habían sido mencionados explícitamente entre los hechos explicativos. Luego, no está claro por qué sólo a

Una clara exposición y una apreciación crítica de las diversas razones que se han aducido en apoyo de la idea de concebir las leyes generales como reglas de inferencia se encontrarán en Alexander (1958).

[31] Scriven (1959, en particular la sección 3.1).

[32] Ryle (1950, pág. 330).

las leyes se les debe atribuir la función justificadora.[33] Y si se admite que los enunciados que expresan hechos particulares también pueden servir como fundamentos justificadores en las explicaciones, entonces la distinción entre hechos explicativos y fundamentos justificadores se hace oscura y arbitraria.

Scriven va más allá de relegar las leyes explicativas al papel de fundamentos justificadores: sostiene que, a veces, podemos estar totalmente seguros de una explicación sin poder justificarla mediante referencias a leyes; según sus propias palabras, "ciertos elementos de juicio bastan para garantizar ciertas explicaciones sin necesidad de la deducción a partir de leyes".[34] Uno de los ejemplos que ofrece es el siguiente:

> Mientras usted trata de alcanzar el diccionario, su rodilla tropieza con el borde de la mesa y vuelca el tintero, cuyo contenido se derrama en el suelo y arruina la alfombra. Si luego se le pide una explicación de cómo se arruinó la alfombra, usted dispone de una explicación completa. Usted la arruinó al volcar la tinta. La certeza de esta explicación es primigenia. No tiene absolutamente nada que ver con su conocimiento de las leyes de la física atinentes al caso; un hombre de las cavernas podría dar la misma explicación y sentirse igualmente seguro de ella... Si se le pidiera a usted que presentara los fundamentos justificadores de su explicación, ¿qué podría hacer? *No podría presentar ninguna hipótesis universal verdadera* en la cual el antecedente estuviera identificablemente presente (es decir, que evitara expresiones tales como "tropezar con bastante fuerza") y el consecuente fuera el efecto que se quier explicar.[35]

A lo sumo, continúa Scriven, se podría ofrecer una vaga generalización según la cual si tropezamos con una mesa con bastante fuerza, hará que un tintero colocado en un lugar no muy seguro se derrame, siempre que haya en él bastante tinta. Pero se la debe precisar de muchas maneras y, sostiene Scriven, no se la puede convertir en una hipótesis universal verdadera que, para el ejemplo en cuestión, "salvara el modelo deductivo". En particular, no cabe esperar que la física brinde tal hipótesis, pues "la explicación no se ha hecho un ápice más segura desde que se descubrieron las leyes de la elasticidad y la inercia".[36]

Indudablemente, en nuestras empresas cotidianas y también en discusiones científicas, a menudo ofrecemos o aceptamos explicaciones como la ilustrada por el ejemplo de Scriven. Pero un estudio analítico de la explicación no puede contentarse con registrar simplemente este hecho: se lo debe tratar como material de análisis; se debe tratar de aclarar qué es lo que se *afirma* mediante un enunciado explicativo de este género y cómo puede *sustentarse* tal afirmación. Y Scriven no ofrece ninguna respuesta explícita, al menos para la primera cuestión. No nos dice exactamente qué

33 La misma observación ha hecho Alexander (1958, sección 1).
34 Scriven (1959, pág. 456).
35 *Loc. cit.*, bastardillas del autor.
36 *Loc cit.*

es, en su opinión, lo que se afirma mediante la explicación ofrecida sin leyes; por lo tanto, queda en la oscuridad cuál es, precisamente, la afirmación a la que atribuye una certeza primigenia, para el hombre de las cavernas y para el físico moderno por igual. Presumiblemente, la explicación que tiene *in mente* se expresaría por un enunciado según el cual, en líneas generales, la alfombra se manchó con tinta porque se tropezó con la mesa. Pero, sin duda, este enunciado afirma por implicación que las circunstancias antecedentes invocadas son de un tipo que, por lo general, produce efectos como el que se quiere explicar. En realidad, es justamente esta afirmación implícita de conexiones uniformes generales lo que distingue la atribución causal hecha aquí de una mera narración según la cual primero se tropezó con la mesa, luego se cayó el tintero y, finalmente, se derramó la tinta en la alfombra. Ahora bien, en un caso como en el del derrame de la tinta, estamos familiarizados, al menos de una manera general, con las conexiones uniformes atinentes al caso, aunque no podamos enunciarlas con precisión, por lo cual propendemos a darlas por supuestas sin mención explícita. Por otra parte, hay varios antecedentes concebidos particulares, cualquiera de los cuales podría explicar el vuelco del tintero, en virtud de las mismas uniformidades generales, aproximadamente: yo podría haber pateado la mesa, el gato podría haber empujado el tintero, la cortina podría haber barrido el tintero movida por una brisa, etc. Así, la cuestión de cómo apareció la mancha de tinta habitualmente tenderá a obtener información acerca de los antecedentes particulares que dieron origen al daño; parecería, pues, que una explicación no tiene nada que ver con uniformidades o leyes. Pero esta aparición no refuta, sin duda, la idea de que toda afirmación explicativa particular hecha en términos de circunstancias antecedentes presupone adecuadas leyes abarcantes.

Esto nos lleva a una cuestión fundamental que plantea el argumento de Scriven. ¿Es posible especificar, en el caso dado, un conjunto de leyes que realmente brinde una justificación, permitiéndonos deducir el explanandum, dada la información acerca de los sucesos explicativos antecedentes? No es posible responder a esta pregunta de manera inequívoca porque es demasiado vaga. Suponiendo que el enunciado explicativo adopte la forma "*q* porque *p*", no se nos ha dicho con precisión, en el caso del tintero volcado, qué es lo que ocupa el lugar de "*p*" y de "*q*". Por ejemplo, si se considerara que el enunciado "*p*" incluye la información de que fue empujado un tintero lleno y sin tapa, y si el enunciado "*q*" informara meramente que la tinta se derramó, entonces ciertas leyes elementales de la mecánica de fluidos bien podrían suministrar una base nomológica adecuada para el enunciado explicativo. En cambio, si se considera que el enunciado "*q*" especifica no sólo que la tinta se derramó, sino también que provocó en la alfombra una mancha de tamaño y forma determinados, entonces, indudablemente, no se conocen leyes que permitan inferir del enunciado "*p*" (en cualquier concepción plausible) este enunciado "*q*". Pero, justamente por esta razón, de una explicación como la sugerida por el ejemplo de Scriven no se consideraría que explique el tamaño o la forma de la mancha de tinta.

Sin duda, la afirmación explicativa que tiene en vista Scriven está

entre esos extremos y diría, aproximadamente, que la alfombra se manchó porque mi rodilla tropezó y ladeó la mesa, en la cual había un tintero abierto. Podríamos parafrasear esta afirmación diciendo que hay leyes que vinculan la presencia de una mancha de tinta sobre la alfombra con ciertas circunstancias antecedentes, entre las que figuran un tintero abierto sobre la mesa y el hecho de que la mesa se ladeó. Y no parece haber razón alguna para dudar de la posibilidad de aducir o establecer un conjunto gradualmente creciente de leyes que suministren explicaciones cada vez más exactas y detalladas del fenómeno en cuestión.

Podríamos decir, en coincidencia con Scriven, que estas leyes dan apoyo o justificación a la respuesta dada de la forma "porque". Pero debemos observar también que el aumento del conjunto de leyes sustentadoras normalmente exigirá un aumento correspondiente del conjunto de circunstancias antecedentes que es menester tomar en consideración y, por ende hablando estrictamente, exige una modificación del mismo enunciado explicativo de la forma "porque".

Además, la tarea de establecer los enunciados, de leyes o de hechos particulares, que puedan invocarse en apoyo de un enunciado de la forma "porque" entra, evidentemente, en el dominio de la investigación científica; por consiguiente, no puede argüirse razonablemente que el progreso en la investigación física o química no tiene ninguna significación para la explicación en cuestión. Así, el hombre de las cavernas de Scriven, o quizás un niño, podrían suponer que cuando se vuelca un líquido opaco sobre cualquier tipo de tejido lo impregna y provoca una mancha; lo cual lo llevaría a esperar que aparezca una mancha cuando se vuelca mercurio sobre una alfombra o cuando se vuelca tinta sobre un tejido especialmente tratado contra las manchas. Y si su explicación o interpretación de la mancha de tinta sobre la alfombra presupone esto, entonces estaría lejos de la certeza primigenia: sería simplemente falso.

En resumen, pues, la afirmación de que el hombre de la cavernas podría explicar la mancha de la alfombra con la misma "certeza" que un científico moderno pierde su sorprendente plausibilidad inicial cuando nos preguntamos qué afirma precisamente la explicación y qué implica, y cuando nos aseguramos de que no se la toma simplemente como una narración de etapas elegidas del proceso en cuestión. Una explicación bien puede expresarse bajo la forma de una narración ordenada, pero sólo explicará si, al menos tácitamente, presupone ciertas conexiones nómicas entre las diferentes etapas citadas. Más adelante examinaremos con mayor detalle estas explicaciones "genéticas".

En la discusión anterior hemos concebido un enunciado explicativo de la forma "*q* porque *p*" como una afirmación según la cual: sucede (o sucedió) *p*, y hay leyes (no especificadas de manera explícita) por las cuales el enunciado de que sucede (o sucedió) *q* se sigue lógicamente de esas leyes junto con el enunciado de *p* y, quizás, otros enunciados que especifican antecedentes no incluidos en *p* pero supuestos tácitamente en la explicación. En su examen del papel explicativo de las leyes, Scriven considera la idea íntimamente relacionada con las anteriores de que cuando podemos especificar la causa de un suceso particular como la mancha de

la alfombra, "estamos en condiciones de juzgar, no que se aplican ciertas leyes especificables, sino que deben aplicarse *algunas leyes*". Y objeta que "es *muy* extraño decir esto, y no más bien que a veces podemos estar totalmente seguros de algunos enunciados causales aun cuando no conozcamos leyes atinentes al caso. Esta capacidad para identificar causas se aprende, está más desarrollada en algunas personas que en otras, puede ser sometida a prueba y es la base de lo que llamamos *juicios*".[37]

Pero esta objeción, sin duda, no es muy sólida. Pues, ante todo, para que la tesis tenga un significado claro, necesitamos saber exactamente qué se entiende por "identificar la causa de un suceso particular" y, por consiguiente, cómo puede someterse a prueba la capacidad para identificar causas; y Scriven no brinda esta información.

En segundo término, la idea de que un enunciado de la forma "*q* porque *p*" afirma, por implicación, la existencia de ciertas leyes abarcantes no es en modo alguno incompatible con la opinión de que las personas puedan tener capacidad para el juicio causal aunque sean incapaces de especificar leyes abarcantes o de explicar la noción de causa que utilizan. Hagamos una analogía. Un carpintero o un jardinero experimentados pueden tener la capacidad de juzgar muy exactamente el tamaño de la región encerrada dentro de un círculo dado sin ser capaces de ofrecer una definición analítica del área de un círculo en función de las series convergentes formadas por las áreas de ciertos polígonos inscriptos o circunscriptos. Pero esto no justifica la afirmación de que al menos en los casos específicos accesibles a los juicios de artesanos hábiles, el análisis matemático del concepto de área de un círculo es ajeno a la cuestión o no se aplica. De igual modo, un médico, un mecánico de garaje o un electricista pueden tener una notable capacidad para discernir las causas de un desperfecto en un caso particular sin ser siempre capaces de aducir leyes generales en apoyo del diagnóstico y hasta sin creer que éste presupone la existencia de tales leyes. Pero esto no garantiza la conclusión de que sea imposible o inadecuado concebir los enunciados causales en cuestión como haciendo referencia a ciertas leyes o, al menos, como implicando su existencia.

Aun la manera de someter a prueba y fundamentar los enunciados causales basados en tales juicios "prácticos" indica que, al menos implícitamente, hacen una afirmación de carácter general. Así, la afirmación de que cierta medida terapéutica causa la mejoría en un caso dado requeriría su corroboración por resultados similares en casos similares, para descartar la posibilidad de una mera coincidencia, en contraposición a una conexión causal.

Pero puesto que a menudo las explicaciones se formulan como enunciados de la forma "porque", ¿no debemos introducir al menos un modelo adicional que conciba las explicaciones como enunciados de la forma "*q* porque *p*", y no como razonamiento? Caracterizar cierto tipo de explicación diciendo simplemente que es de esta forma, sería insuficiente, sin duda: la tarea principal del modelo considerado sería clarificar el significado de la palabra "porque" en contextos explicativos, y esto exige mayor análisis. Afirmar que podemos a veces dar explicaciones de la forma "*q*

[37] *Loc. cit.*, bastardillas del autor.

porque p" con completa certeza o que se las puede fundamentar mediante adecuados elementos de juicio prescindiendo de leyes, es soslayar este problema; en realidad, tal afirmación ni siquiera puede evaluarse independientemente de un análisis del uso explicativo de la palabra "porque". La paráfrasis de los enunciados de la forma "porque" sugerida antes es más bien vaga y, sin duda, se la puede mejorar, pero al menos me parece que pone de manifiesto correctamente la suposición de conexiones legales implícitas en tales formulaciones explicativas.

2.4. *El carácter potencialmente predictivo de la explicación*

Debido a su esencial dependencia de leyes y principios teóricos, cabe esperar que la explicación D-N presente una estrecha afinidad con la predicción científica; pues las leyes y los principios teóricos, al hacer afirmaciones generales, abarcan también casos aún no examinados y tienen implicaciones definidas para ellos.

Ilustra vívidamente esta afinidad la parte cuarta de los *Diálogos sobre dos nuevas ciencias.* En esta obra, Galileo desarrolla sus leyes sobre el movimiento de los proyectiles y deduce de ellas el corolario de que si se disparar proyectiles desde el mismo punto con igual velocidad inicial pero diferentes elevaciones, se obtendrá el alcance máximo cuando la elevación sea de 45°. Luego, Galileo hace observar a Sagredo: "A través de los relatos de artilleros, ya conocía el hecho de que en el uso de cañones y morteros el alcance máximo ... se obtiene cuando la elevación es de 45° ...; pero la comprensión de por qué esto ocurre supera en mucho la mera información obtenida por el testimonio de otros o siquiera por la repetida experimentación".[38] El razonamiento que brinda tal comprensión puede expresarse fácilmente en la forma (D-N); equivale a una deducción, por medios lógicos y matemáticos, del corolario a partir de un conjunto de premisas que contiene: (I) las leyes fundamentales de la teoría de Galileo sobre el movimiento de los proyectiles, y (II) enunciados particulares en los cuales se especifica que todos los proyectiles considerados son disparados desde el mismo lugar con la misma velocidad inicial. Evidentemente, pues, se *explica* y, por ende, se *comprende* el fenómeno observado anteriormente por los artilleros mostrando que es de esperar que se produzca, en las circunstancias especificadas, en vista de ciertas leyes generales expuestas en la teoría de Galileo. Y el mismo Galileo señala con obvio orgullo las *predicciones* que pueden hacerse, de igual manera, mediante deducciones a partir de sus leyes; éstas implican "lo que nunca se ha observado en la experiencia, a saber, que los tiros que se exceden o no llegan a los 45° en cantidades iguales tienen alcances iguales". Así, la explicación que suministra la teoría de Galileo "prepara a la mente para comprender y discernir otros hechos sin necesidad de recurrir a experimentos",[39] es decir, por subsunción deductiva bajo las leyes en las que se basa la explicación.

[38] Galilei (1946, pág. 265).
[39] *Loc. cit.*

Controlar las predicciones derivadas de este modo de las leyes generales o los principios teóricos invocados en una explicación es una manera importante de poner a prueba esas generalizaciones "abarcantes", y un resultado favorable puede prestarles un firme apoyo. Consideremos, como ejemplo, la explicación ofrecida por Torricelli de un hecho que había intrigado a su maestro Galileo, a saber, que una bomba elevadora de agua de un pozo no levanta el agua a más de 10 metros por encima de la superficie del pozo.[40] Para explicar esto, Torricelli expuso la idea de que el aire que está sobre el agua tiene peso y, por consiguiente, ejerce presión sobre el agua del pozo, obligándola a elevarse por el caño de la bomba cuando se levanta el pistón, pues no hay aire dentro de él que equilibre la presión externa. Según esta suposición, el agua sólo puede elevarse hasta el punto en el cual su presión sobre la superficie del pozo es igual a la presión del aire externo sobre esta superficie, la cual será igual, pues, a la de una columna de agua de unos 10 metros de altura.

La fuerza explicativa de esta descripción depende de la concepción según la cual la Tierra está rodeada por un "mar de aire" que satisface las leyes básicas que gobiernan el equilibrio de líquidos en vasos comunicantes. Y puesto que la explicación de Torricelli presuponía tales leyes generales, brindaba predicciones concernientes a fenómenos aún no examinados. Una de éstas era que si se reemplazaba el agua por mercurio, cuyo peso específico es unas 14 veces mayor que el agua, el aire equilibraría una columna de unos 10/14 metros, o sea, más o menos unos 76 cm de longitud. Esta predicción fue confirmada por Toricelli en el experimento clásico que lleva su nombre. Además, la explicación propuesta implica que a medida que aumenta la altura por encima del nivel del mar, debe disminuir la longitud de la columna de mercurio soportada por la presión del aire, porque disminuye el peso del aire que la equilibra. Se realizó una prueba cuidadosa de esta predicción, a sugerencia de Pascal, sólo pocos años después de que Torricelli ofreciera su explicación: el cuñado de Pascal llevó un barómetro de mercurio (es decir, en esencia, una columna de mercurio equilibrada por la presión del aire) hasta la cima del Puy-de-Dôme, y midió la longitud de la columna a diversas alturas durante el ascenso y durante el descenso; los resultados concordaban muy bien con la predicción.[41]

Las inferencias mediante las cuales se obtienen tales predicciones son, nuevamente, de forma deductivo-nomológica: las premisas comprenden las leyes explicativas en cuestión (en nuestro último ejemplo, la hipótesis de Torricelli especialmente) y ciertos enunciados de hechos particulares (por ejemplo, que se llevará a la cima de una montaña un barómetro de tales y cuales características). Llamemos predicciones D-N a los razonamientos predictivos de la forma (D-N). En la ciencia empírica, muchos razonamientos predictivos son de este tipo. Entre los ejemplos más notables se cuentan las previsiones, basadas en los principios de la mecánica celeste y

[40] La siguiente explicación se basa en la exposición de este caso que hace Conant (1951, capítulo 4).

[41] La propia explicación de Pascal y su apreciación del "gran experimento" están reimpresas en la traducción inglesa de Moulton y Schifferes (1945, pág. 144-153).

de la óptica, concernientes a las posiciones relativas del Sol, la Luna y los planetas en un momento dado, y concernientes a eclipses solares y lunares.

Puede ser conveniente destacar aquí que si bien los principios de la mecánica celeste y otras leyes o teorías deterministas suministran la base para realizar explicaciones y predicciones D-N muy impresionantes, las premisas adicionales que se requieren para este propósito no sólo deben brindar la especificación del estado del sistema en un tiempo t_0 anterior al tiempo t_1 para el cual debe inferirse el estado del sistema, sino también un enunciado sobre las condiciones limitadas que prevalecen entre t_0 y t_1; éstas especifican las influencias externas que se ejercen sobre el sistema durante el intervalo de tiempo en cuestión. Para ciertos propósitos astronómicos puede despreciarse la influencia perturbadora de otros objetos celestes, aparte de los considerados explícitamente, y considerar el sistema como "aislado"; pero esto no debe hacernos pasar por alto el hecho de que ni siquiera en esos casos de predicción deductivo-nomológica podemos prever sucesos futuros estrictamente sobre la base de la información acerca del presente; el razonamiento predictivo también exige ciertas premisas concernientes al futuro, por ejemplo, la ausencia de influencias perturbadoras, como un choque de Marte con un cometa inesperado; y el alcance temporal de estas condiciones limitantes debe extenderse hasta el tiempo mismo de la aparición del suceso predicho. Por lo tanto, la afirmación de que las leyes y teorías de forma determinista nos permiten predecir ciertos aspectos del futuro a partir de la información acerca del presente debe ser tomada con moderación. Observaciones análogas se aplican a la explicación deductivo-nomológica.

Puesto que en una explicación D-N totalmente explícita de un hecho particular el explanans implica lógicamente el explanandum, podemos decir que el razonamiento explicativo podría haber sido usado para una predicción deductiva del hecho del explanandum *si* las leyes y los hechos particulares en su explanans hubieran sido conocidos y tomados en consideración en un tiempo anterior adecuado. En este sentido, una explicación D-N es una predicción D-N potencial.

Oppenheim y yo hicimos ya esta observación en un artículo anterior[42] en el cual añadimos que la explicación científica (del tipo deductivo-nomológico) no difiere de la predicción científica por su estructura lógica, sino por ciertos aspectos pragmáticos. En un caso, se sabe que se ha producido el suceso descripto en la conclusión, y se buscan enunciados adecuados que expresen leyes generales y hechos particulares para explicarlo; en el otro, se dispone ya de estos enunciados, y de ellos se deduce el correspondiente al suceso en cuestión antes del momento de su presunta aparición. Esta concepción, que ha sido llamada a veces la tesis *de la identidad estructural* (o de la simetría) *de la explicación y de la predicción*, ha sido puesta en tela de juicio por varios autores. El examen de algunos de sus argumentos puede ayudarnos a aclarar el problema implicado.

[42] Hempel y Oppenheim (1948, sección 3).

Para comenzar, algunos autores[43] han observado que lo que se llama por lo común una predicción no es un razonamiento sino una proposición. Más precisamente, como ha señalado Scheffler, es un signo de proposición, es decir, la pronunciación o inscripción concreta de una proposición que se propone describir un hecho que debe suceder después de la producción del signo.[44] Esto es así, por cierto. Pero en la ciencia empírica, normalmente se establecen las proposiciones predictivas sobre la base de la información disponible por medio de razonamientos que pueden tener un carácter deductivo o inductivo; y la tesis en discusión debe entenderse, por supuesto, como refiriéndose a *razonamientos* explicativos y predictivos.

Concebida de este modo, *la tesis de la identidad estructural equivale a* la conjunción de *dos subtesis,* a saber: 1) que *toda explicación adecuada es potencialmente una predicción,* en el sentido ya indicado; 2) que, a la inversa, *toda predicción adecuada es potencialmente una explicación.* Examinaré ahora una serie de objeciones que se han planteado respecto de dicha tesis, comenzando con aquellas que, en efecto, se relacionan con la primera subtesis, para tratar luego las concernientes a la segunda. Sostendré que la primera subtesis es válida, mientras que la segunda está sujeta a dudas. Aunque las consideraciones siguientes se refieren principalmente a la explicación D-N, algunas de ellas son también aplicables a otros tipos de explicación. La adecuación de la tesis de la identidad estructural al caso de la explicación estadística será examinada con detalle en la sección 3.5.

La primera subtesis, como ya se ha observado, es una verdad casi trivial en el caso de la explicación D-N, ya que en éste el explanans implica lógicamente el explanandum. Pero también recibe apoyo de un principio más general que se aplica igualmente a otros tipos de explicación y que expresa, creo, una *condición general de adecuación para toda explicación racionalmente aceptable de un hecho particular.* Esta condición es la siguiente: toda respuesta racionalmente aceptable a la pregunta "¿por qué se produjo el hecho *X*?" debe ofrecer una información que muestre que *X* era de esperar, si no definidamente como en el caso de la explicación D-N, al menos con razonable probabilidad. Así, la información explicativa debe brindar buenos fundamentos para creer que *X*, de hecho, ocurrió. En caso contrario, esta información no nos brindaría ninguna razón adecuada para afirmar: "Esto lo explica; esto muestra por qué sucedió *X*". Y una descripción explicativa que satisfaga esta condición constituye, por supuesto, una predicción potencial en el sentido de que hubiera servido para predecir la producción de *X* (deductivamente o con una probabilidad más o menos alta), si se hubiera tenido la información contenida en el explanans en algún momento anterior.

Puede extenderse, de una manera obvia, la condición de adecuación que acabamos de enunciar a explicaciones referentes, no a hechos individuales, sino a uniformidades empíricas expresadas por leyes putativas.

[43] Véase Scheffer (1957, sección 1, y 1963, Parte I, secciones 3 y 4) y Scriven (1962, pág. 177).

[44] Véase Scheffler (1957, sección 1). Para un estudio más detallado de la explicación y la predicción a la luz de la distinción entre tipo y signo, véase Kim (1962).

Pero no puede decirse de tales explicaciones que son *predicciones* potenciales, ya que los enunciados legales se proponen expresar uniformidades intemporales y, por ende, no hacen referencia alguna a un tiempo particular pasado, presente o futuro.[45]

Casi no es necesario destacar que el *propósito* de una explicación no es, por supuesto, dar razones en apoyo del enunciado-explanandum; pues, como observamos en la primera sección de este ensayo, la búsqueda de una explicación normalmente *presupone* que el enunciado-explanandum es verdadero. El objeto de las observaciones precedentes es, más bien, señalar que una explicación adecuada no puede menos que suministrar información que, si se la establece adecuadamente, también brinda razones en apoyo del enunciado-explanandum. En la terminología de la sección 1, podemos decir que una respuesta adecuada a una pregunta sobre el porqué que pide una explicación también es siempre una respuesta potencial a la correspondiente pregunta sobre el porqué epistémico.

Pero la afirmación inversa no es válida; la condición de adecuación es necesaria pero no suficiente para que una explicación sea aceptable. Por ejemplo, ciertos hallazgos empíricos pueden suministrar excelentes razones en apoyo de la creencia de que la orientación del campo magnético terrestre presenta variaciones diurnas y seculares, sin explicar en absoluto por qué sucede esto. De manera similar, un conjunto de datos experimentales puede *apoyar* firmemente la suposición de que la resistencia eléctrica de los metales aumenta con su temperatura o que ciertos productos químicos impiden el crecimiento de células cancerosas, sin suministrar ninguna *explicación* de estas presuntas regularidades empíricas. Las inferencias predictivas aquí implicadas son inductivas más que deductivas; pero lo que impide que adquieran el status de explicaciones potenciales no es su carácter inductivo (en la sección 3 trataremos de razonamientos inductivos que brindan explicaciones científicas absolutamente satisfactorias), sino el hecho de que no invocan leyes o principios teóricos, es decir, enunciados explicativos que hagan afirmaciones generales. La apelación a principios relacionantes generales, si bien no son indispensables para la predicción, se requieren para toda explicación: solamente tales principios pueden dar a cualquier circunstancia particular que pueda aducirse el *status* de factores explicativos del hecho en cuestión.

Algunas de las objeciones planteadas acerca de las tesis de la identidad estructural de la explicación y la predicción se refieren a la primera de sus dos subtesis, que hemos expuesto con algún detalle: es decir, a la afirmación de que todo razonamiento explicativo adecuado es también potencialmente predictivo. Examinaré tres objeciones según las cuales hay ciertas explicaciones totalmente satisfactorias que no constituyen predicciones potenciales.

Scriven ha sostenido que, a veces, un hecho *X* queda adecuadamente explicado mediante una "proposición de la forma 'la única causa de *X* es *A*...', por ejemplo, 'la única causa de la paresia es la sífilis' "; esta proposición nos permite explicar por qué determinado paciente tiene paresia señalando que había sufrido de sífilis. Y esta explicación es válida, según

[45] Destaca este punto, por ejemplo, Scriven (1962, págs. 179 y sigs.).

Scriven, aunque sólo un pequeño porcentaje de pacientes sifilíticos la manifiesten, de modo que "en presencia de los elementos de juicio [de que una persona determinada tiene sífilis], aún debamos hacer la predicción de que [la paresia] no surgirá".[46] Pero si aparece, el principio de que la única causa de la paresia es la sífilis puede "brindar y garantizar nuestra explicación" en función de la infección sifilítica antecedente.[47] Así, tenemos aquí una presunta explicación que, en verdad, no es adecuada como predicción potencial. Pero precisamente porque la paresia es una secuela tan rara de la sífilis, la anterior infección sifilítica no puede, por sí sola, suministrar una adecuada explicación de ella. Una condición que es nómicamente necesaria para la producción de un hecho, en general no lo explica; de otro modo, podríamos explicar que un hombre ganó el primer premio en la combinación irlandesa en las carreras señalando que previamente había comprado un billete y que sólo una persona que posee un billete puede ganar el primer premio.

Un segundo argumento que, como el de Scriven, tiene considerable plausibilidad inicial ha sido expuesto por Toulmin[48] con referencia a la "teoría de Darwin, que explica el origen de las especies por la variación y la selección natural. Ningún científico ha utilizado nunca esta teoría para predecir la aparición de seres de una nueva especie, y menos aún ha verificado su predicción. Sin embargo, muchos científicos competentes han aceptado la teoría de Darwin por su gran poder explicativo". Al examinar este argumento, permítaseme distinguir entre lo que podría llamarse la *historia* de la evolución y la *teoría* de los mecanismos subyacentes de la mutación y la selección natural. La historia de la evolución, como hipótesis acerca del desarrollo gradual de varios tipos de organismos y acerca de la ulterior extinción de muchos de éstos, tiene el carácter de una narración histórica hipotética que *describe* las presuntas etapas del proceso

[46] Scriven (1959a, pág. 480); bastardillas del autor.

[47] *Loc. cit.* Barker ha sostenido, análogamente, que "puede ser correcto hablar de explicación en muchos casos en que no es posible realizar una predicción específica. Así, por ejemplo, si el paciente muestra todos los síntomas de la neumonía, se enferma y muere, puedo entonces explicar su muerte, pues sé qué es lo que lo mató, pero no podía de antemano haber predicho definitivamente la muerte, pues habitualmente la neumonía no es fatal" (1961, pág. 271). Me parece que este argumento está sujeto a objeciones similares a las planteadas con referencia al ejemplo de Scriven. Ante todo, no está claro qué es lo que se afirma al decir que la neumonía mató al paciente. Sin duda, la mera información de que el paciente tenía neumonía no basta para explicar su muerte, precisamente porque en la mayoría de los casos no es fatal. Y si suponemos que el explanans afirmaba que el paciente padecía de una neumonía muy grave (y quizá que era viejo y débil), entonces bien podía brindar una base al menos para una explicación probabilística de la muerte del enfermo; pero en este caso, es obvio que también permitiría una predicción de su muerte con la misma probabilidad. Otras observaciones sobre el argumento de Barker se encontrarán en los comentarios de Feyerabend y Rudner, y las réplicas de Barker, publicados en Feigl y Maxwell (1961, págs. 278-285). Una detallada discusión crítica, que arroja nueva luz sobre el ejemplo de la paresia que da Scriven, se hallará en Grünbaum (1963 y 1963a, cap. 9); véase también la respuesta de Scriven (1963).

[48] Toulmin (1961, págs. 24-25), Scriven (1959a) y Barker (1961) han expuesto argumentos similares. Para una discusión crítica de la versión de Scriven, véase Grünbaum (1963 y 1963a, cap. 9).

evolutivo; pero es la teoría asociada a ella la que suministra la *visión explicativa* que poseemos de este proceso. La historia de la evolución puede decirnos, por ejemplo, que en cierta etapa del proceso hicieron su aparición los dinosaurios y que, mucho más tarde, se extinguieron. Tal descripción, por supuesto, no explica por qué surgieron los diversos tipos de dinosaurios con sus características distintivas, ni explica por qué se extinguieron. En realidad, ni siquiera la teoría de la mutación y la selección natural asociada a esta descripción responde a la primera de estas cuestiones, aunque podría sostenerse que aclara un poco la segunda. Sin embargo, aun para explicar la extinción de los dinosaurios necesitamos una gran cantidad de hipótesis adicionales acerca de su medio físico y biológico, y de las especies con las que tuvieron que competir por la supervivencia. Pero si poseemos hipótesis de este tipo que sean suficientemente específicas como para suministrar, en combinación con la teoría de la selección natural, al menos una explicación probabilística de la extinción de los dinosaurios, entonces el explanans aducido también puede ser la base de una potencial predicción probabilística. La persuasividad innegablemente grande del argumento de Toulmin parecería derivar de dos fuentes: una difundida tendencia a creer que la historia, básicamente descriptiva, de la evolución, explica los diversos estadios del proceso, y una tendencia de igual modo difundida a sobreestimar la medida en la cual aun la teoría de la mutación y de la selección natural puede explicar los detalles de la sucesión evolutiva.

Pasaré ahora a una tercera objeción respecto de la afirmación de que una explicación adecuada es también una predicción potencial. Se basa en la observación de que, a veces, la única base que tenemos para afirmar algún enunciado esencial del explanans reside en el conocimiento de que el hecho del explanandum realmente ocurrió. En tales casos, es evidente que el argumento explicativo no podría haberse usado para predecir este hecho. Consideremos uno de los ejemplos de Scriven.[49] Supongamos que un hombre ha dado muerte a su mujer de quien sabía que le era infiel, y que se explica esta acción como el resultado de intensos celos. El hecho de que el hombre sintiera celos podría haberse previsto antes del suceso, pero para explicar este último necesitábamos saber que sus celos eran bastante intensos como para llevarlo a cometer el crimen; y esto sólo podemos saberlo después de haberse cometido aquél. En este caso, pues, la producción del hecho del explanandum suministra el único fundamento que tenemos para afirmar una parte importante del explanans; por lo tanto, el suceso del explanandum no podía haberse predicho por medio del razonamiento explicativo. En otro ejemplo[50] Scriven examina una explicación según la cual el derrumbe de un puente fue causado por la fatiga de los metales. Podría apoyarse esta explicación, arguye, señalando que el derrumbe sólo podía haberlo causado una carga excesiva, por daño externo o por fatiga de los metales, y que los dos primeros factores no estuvieron presentes en el caso en cuestión, mientras que hay pruebas de fatiga de los metales. *Dada la información de que el puente, de hecho, se derrumbó,*

[49] Scriven (1959, págs. 468-469).
[50] Scriven (1962, págs. 181-187).

esto demostraría no sólo que había fatiga en los metales, sino también que era bastante intensa como para provocar el derrumbe. Si bien la noción de Scriven de "la única causa posible" de un suceso dado requiere indudablemente mayor elucidación, su ejemplo ilustra una descripción explicativa tal que una de sus hipótesis constituyentes sólo se sustenta en la producción del suceso que se quiere explicar, de modo que éste no podía haberse predicho por medio del argumento explicativo.

Sin embargo, el punto ilustrado de este modo no afecta en absoluto a la tesis condicional de que un argumento explicativo adecuado debe ser tal que pueda servir para predecir el suceso del explanandum *si* se conociera y se tomara en consideración la información incluida en el explanans antes de que se produjera el hecho. Los casos que presenta Scriven sólo muestran que a veces no sabemos, independientemente de la producción del hecho del explanandum, si se han dado todas las condiciones enumeradas en el explanans. Pero esto sólo significa que, en tales casos, nuestra tesis condicional es contrafáctica, es decir, que su cláusula antecedente no se satisface, pero no que la tesis misma sea falsa. Además, el argumento de Scriven ni siquiera muestra que en el tipo de casos que menciona sea lógica o nomológicamente imposible (imposible en razón de las leyes de la lógica o de las leyes de la naturaleza) para nosotros conocer el factor explicativo crítico antes o independientemente de la aparición del suceso del explanandum; la imposibilidad parece ser, más bien, de carácter práctico y quizá temporario, en reflejo de limitaciones actuales del conocimiento o la tecnología.

Pero aunque no afecte a nuestra tesis, la observación de Scriven presenta interés metodológico en sí misma: muestra que, a veces, se explica un suceso mediante hipótesis cuyo único sustento disponible es el hecho de que se haya producido. Esto puede suceder, como vimos, cuando una de las hipótesis explicativas declara que cierto factor importante fue suficientemente fuerte como para provocar el hecho en cuestión; pero la observación también se aplica a otros casos. Así, la explicación esbozada en la sección 2.1 de la aparición y crecimiento inicial de las pompas de jabón incluía en su explanans la suposición de que se había formado una película de jabón entre la plancha y el borde de los vasos; y prácticamente la única prueba disponible en apoyo de esta suposición explicativa era el hecho de que surgieran pompas de jabón debajo de los vasos. O examinemos la explicación de las líneas oscuras características del espectro de absorción de una estrella particular. La suposición fundamental del explanans es que la atmósfera de la estrella contiene ciertos elementos, tales como hidrógeno, helio y calcio, cuyos átomos absorben radiación de las longitudes de onda correspondientes a las líneas oscuras; la explicación se basa, por supuesto, en muchas otras suposiciones entre las que se cuentan las de la teoría óptica que constituye la base de la espectroscopia, y la suposición de que el aparato utilizado es un espectroscopio bien construido. Pero, si bien estos últimos enunciados del explanans pueden recibir prueba y corrobación independientes, bien puede suceder que el único elemento de juicio en apoyo de la hipótesis explicativa fundamental sea la aparición de las mismas líneas cuya presencia en el espectro pretende ex-

plicar el argumento. Hablando en términos estrictos, el suceso del explanandum brinda apoyo a la hipótesis explicativa fundamental sólo en virtud de la teoría en la cual se respalda, que vincula la presencia de ciertos elementos en la atmósfera de una estrella con la aparición en su centro de líneas de absorción correspondientes. Así, la información de que se ha producido el suceso del explanandum no da apoyo, por sí mismo, a la hipótesis explicativa en cuestión, pero constituye, podríamos decir, una parte esencial de los únicos elementos de juicio disponibles en apoyo de tal hipótesis.

Las explicaciones del tipo considerado pueden caracterizarse esquemáticamente como razonamientos de la forma (D-N), en los cuales la información o la suposición de que E es verdadero suministra una parte indispensable del único apoyo probatorio disponible para uno de los enunciados del explanans, por ejemplo, C_1. Llamemos a tales explicaciones *autoevidentes*. Podría sostenerse que la producción real del hecho del explanandum suministra siempre un leve apoyo adicional aun a un explanans cuyas proposiciones constituyentes se han aceptado sobre la base de elementos de juicio independientes y que, en este sentido, toda explicación D-N que tenga un explanandum verdadero es, en cierta medida, autoevidente, pero aplicaremos este recurso a una descripción explicativa sólo si, en el momento de su presentación, la producción del suceso del explanandum suministra la única prueba, o una parte indispensable de la única prueba, disponible en apoyo de algunos de los enunciados del explanans. Un razonamiento explicativo de la forma (D-N) que sea autoevidente no es por esta razón circular o inútil. Sin duda, si se adujera el mismo razonamiento en apoyo de la afirmación de que el hecho del explanandum se produjo (o que E es verdadero), entonces sería vulnerable a la acusación de circularidad epistémica. Para que el razonamiento logre su objetivo, pues, todas las razones que aduce en apoyo de E —es decir, C_1, C_2, ..., C_k; L_1, L_2, ..., L_r— deberían establecerse independientemente de E; pero aquí se viola esta condición, ya que la única razón que tenemos para creer o afirmar C_1 incluye la suposición de que E es verdadero. Pero cuando se usa el mismo razonamiento con propósitos explicativos no pretende establecer que E sea verdadero; esto lo presupone la pregunta "¿Por qué se produjo el suceso descripto por E?" Una explicación autoevidente tampoco supone necesariamente un círculo vicioso explicativo. La información de que se ha producido el suceso del explanandum no está incluida en el explanans (de modo que la producción del suceso no "se explica por sí misma"); más bien, totalmente aparte del contexto explicativo, sirve como prueba en apoyo de uno de los enunciados del explanans. Así, una explicación autoevidente aceptable se beneficia, por así decir, con la sabiduría de la visión retrospectiva derivada de la información de que el hecho del explanandum se ha producido, pero no utiliza mal esta información para elaborar una explicación circular.

Una explicación autoevidente, por esta razón, puede descansar en un explanans que tiene poco sustento y por lo tanto, puede no tener firmes bases empíricas. Pero ni siquiera esto es inevitable. En el caso del espectro de absorción de un estrella, por ejemplo, la información básica aceptada

previamente, incluyendo las teorías atinentes al caso, puede indicar que las líneas oscuras observadas *sólo* aparecen si los elementos específicados están presentes en la atmósfera de la estrella; y entonces el explanandum, en conjunción con la información básica, presta muy firme apoyo a la hipótesis explicativa fundamental.

Creo que la noción de explicación autoevidente puede contribuir a aclarar el desconcertante problema ilustrado por la explicación de la paresia en función de una infección sifilítica previa. Tomemos otro ejemplo. Se atribuyen algunos casos de cáncer de piel a la irradiación ultravioleta intensa. Pero muy a menudo este factor no provoca cáncer, de modo que la información de que una persona ha estado expuesta a tal radiación no permite predecir el cáncer. Pero, ¿es suficiente esta única información para explicar la aparición de cáncer de piel, posterior a una intensa irradiación ultravioleta? Sin duda, a menudo se formulará una explicación que sólo mencione la irradiación anterior; pero la justificación subyacente debe ser más compleja. Dejando de lado los importantes aspectos cuantitativos del problema, el quid de esta justificación, creo, puede enunciarse esquemáticamente del siguiente modo: ciertos individuos, aunque en modo alguno todos, tienen disposición a desarrollar cáncer de piel al exponerse a una intensa iradiación ultravioleta; a esta disposición llamémosla sensiblidad a la radiación. Ahora bien, para los propósitos de la explicación, sabemos que un individuo determinado estuvo expuesto a una intensa radiación (C_1) y adquirió cáncer de piel en la región afectada (E). Pero, conjuntamente, estas dos informaciones dan apoyo a la suposición de que el individuo es sensible a la radiación (C_2), hipótesis que no tiene apoyo para los propósitos de la predicción cuando se dispone de C_1, pero no de E. Y los dos enunciados, C_1 y C_2 (en combinación con el enunciado general de que los individuos sensibles adquieren cáncer de piel cuando están expuestos a una radiación intensa), suministran un adecuado explanans para E. De este modo, si se admite que la explicación aduzca C_2 además de C_1, se ve que es autoevidente, pero también que posee un explanans que suministraría una base adecuada para la predicción si C_2 pudiera conocerse de antemano. Esto es imposible, por supuesto, en tanto la única prueba disponible para establecer la sensibilidad a la radiación consista en verificar si un individuo adquiere el cáncer de piel al ser sometido a una intensa irradiación. Pero, evidentemente, es concebible que puedan encontrarse otras pruebas independientes de sensibilidad a la radiación, en cuyo caso C_2 podría ser establecida independientemente, y aun antes, de la producción del suceso descripto por E.

Al examinar la identidad estructural de la explicación y la predicción, sólo he considerado hasta ahora la primera de las dos subtesis mencionadas antes, a saber, la de que toda explicación adecuada es también una predicción potencial. He sostenido que las objeciones planteadas contra esta afirmación no corresponden y que la primera subtesis es correcta y puede servir, en verdad, como condición necesaria de adecuación de toda explicación explícitamente enunciada que sea aceptable de manera racional.

Paso ahora a la segunda subtesis, o sea la de que todo razonamiento predictivo adecuado suministra también una explicación potencial. Esta

afirmación es vulnerable a la duda aun en el caso de ciertos razonamientos predictivos de carácter deductivo-nomológico, como ilustra el siguiente ejemplo. Un de los primeros síntomas del sarampión es la aparición de pequeñas manchas blancuzcas, llamadas manchas de Koplik, en las mucosas de las mejillas. El enunciado *L* según el cual la aparición de manchas de Koplik va siempre seguida por las manifestaciones posteriores del sarampión, pues, podría tomarse como una ley y usarse como premisa en los razonamientos D-N, con una segunda premisa de la forma "el paciente *i* tiene manchas de Koplik en el tiempo *t*", y con una conclusión que declare que posteriormente *i* presenta las manifestaciones ulteriores del sarampión. Un razonamiento de este tipo es adecuado para propósitos predictivos, pero podría ponerse en duda su adecuación como explicación. No queremos decir, por ejemplo, que *i* tuvo fiebre alta y otros síntomas del sarampión porque previamente había tenido manchas de Koplik. Pero este caso, y otros similares a él, no constituye una objeción decisiva contra la segunda subtesis. Pues la renuencia a considerar la aparición de manchas de Koplik como explicación bien puede reflejar dudas acerca de si esas manchas, como ley universal, van siempre seguidas por las ulteriores manifestaciones del sarampión. Quizás una inoculación local con una pequeña cantidad de virus del sarampión produciría las manchas sin provocar la eclosión total de la enfermedad. Si esto fuera así, la aparición de las manchas seguiría suministrando una base habitualmente confiable para predecir la aparición de otros síntomas, ya que las condiciones excepcionales del tipo que acabamos de mencionar serían sumamente raras; pero la generalización de que las manchas de Koplik van siempre seguidas por síntomas posteriores del sarampión no expresaría una ley y, por ende, no podría brindar apoyo adecuado a la correspondiente explicación D-N.

Esta objeción se relaciona con la fuerza explicativa de razonamientos predictivos de la forma (D-N). Pero la segunda subtesis, en su forma general, que no se limita a las predicciones D-N, también ha sido puesta en duda, particularmente por Scheffler y por Scriven,[51] sobre la base de que hay otros tipos de razonamiento predictivo adecuados para la predicción científica, pero no para la explicación. Específicamente, como observa Scheffler, una predicción científica puede basarse en un conjunto finito de datos que no contenga ninguna ley y sin fuerza explicativa alguna. Por ejemplo, un conjunto finito de datos obtenidos en muchas pruebas de ensayo de la hipótesis de que la resistencia eléctrica de los metales aumenta con la temperatura puede brindar un buen apoyo a esta hipótesis y, así, suministrar una base aceptable para la predicción de que, en un caso todavía no examinado, el aumento de temperatura de un conductor metálico irá acompañado por un aumento de la resistencia. Pero si este hecho luego se produce realmente, los datos de la prueba de ensayo evidentemente no suministran una explicación de él. De manera análoga, una lista de los resultados obtenidos en una larga serie de lanzamientos de una moneda determinada puede suministrar una buena base para predecir el porcentaje de caras y cruces que cabe esperar en los mil lanzamientos siguientes de

[51] Véase Scheffler (1957, pág. 296, y 1963, pág. 42); Scriven (1959a, pág. 480).

la misma moneda; pero esta lista de datos tampoco brinda explicación alguna de los resultados posteriores. Los casos como éste plantean la cuestión acerca de si no habrá modos correctos de predicción científica que pasen de hechos particulares a hechos particulares sin necesidad de leyes generales, como parece exigirlas toda explicación adecuada. Ahora bien, los razonamientos predictivos que acabamos de considerar no tienen carácter deductivo, sino probabilístico; y el papel de la inferencia probabilística en la explicación y la predicción será considerado con mayor detalle en la sección 3 de este capítulo. Pero con respecto a la segunda subtesis de la doctrina de la identidad estructural, observemos aquí lo siguiente: las predicciones de nuestros ejemplos pasan de una muestra observada de una población a otra muestra aún no observada; y en algunas teorías actuales de la inferencia probabilística tales razonamientos no dependen de leyes empíricas generales. Según la teoría de la lógica inductiva de Carnap,[52] por ejemplo, tales inferencias son posibles sobre bases puramente lógicas; la información acerca de la muestra dada confiere una probabilidad lógica definida a toda predicción propuesta concerniente a una muestra aún no observada. Por otra parte, ciertas teorías estadísticas de la inferencia probabilística evitan la noción de probabilidades puramente lógicas y califican las predicciones del tipo considerado aquí como correctas sólo en la suposición adicional de que la selección de casos individuales de la población total tenga el carácter de un experimento de azar con ciertas características estadísticas generales. Pero esta suposición, cuando se la enuncia explícitamente, tiene la forma de una ley general de forma estadístico-probabilística; por ello, a fin de cuentas, se realizan las predicciones por medio de leyes abarcantes. Y aunque estas leyes no tienen el carácter estrictamente universal de las invocadas en las explicaciones y predicciones D-N, también pueden desempeñar una función explicativa. Concebidas de tal modo, hasta las predicciones en discusión resultan ser explicaciones potenciales (formuladas de manera incompleta).

Las cuestiones básicas que se discuten en estas diferentes concepciones de la inferencia probabilística son aún objeto de debate y de investigación, y este volumen no es el lugar adecuado para realizar una evaluación más completa de las concepciones opuestas. Por lo tanto, aquí consideraremos como no resueltos problemas concernientes a la segunda subtesis de la doctrina que proclama la identidad estructural de la explicación y la predicción.

3. *La explicación estadística*

3.1. *Leyes de forma estadística*

Dirigimos ahora nuestra atención a las explicaciones basadas en enunciados nomológicos de un tipo que no hemos considerado hasta ahora y que han desempeñado un papel de creciente importancia en la ciencia

[52] Carnap (1950, sección 110).

empírica. Los llamaremos *leyes o principios teóricos de forma estadístico-probabilística* o, para abreviar, *leyes estadísticas.*

La mayor parte de nuestro examen se ocupará del uso explicativo de leyes estadísticas de un tipo muy simple; las llamaremos *leyes de forma estadística básica.* Se trata de enunciados que afirman que la probabilidad estadística de que un hecho de tipo F sea también de tipo G es r, o sea, que, abreviando:

$$p(G,F) = r$$

Hablando en términos generales, este enunciado afirma que, a la larga, la proporción de los casos de F que sean también casos de G es aproximadamente r. (En la sección 3.3 daremos una explicación más completa.)

Por ejemplo, el enunciado de que al rodar un dado ligeramente irregular (hecho de tipo F) salga un as (hecho de tipo G) con una probabilidad de 0,15 —es decir, alrededor del 15 % de todos los casos— tiene esta forma estadística básica. Lo mismo sucede con la ley de que la vida media del radón es de 3,82 días, es decir que la probabilidad estadística de que un átomo de radón se desintegre en cualquier período dado de 3,82 días es de ½, lo cual significa, aproximadamente, que en una muestra de radón que contenga un número grande de átomos, muy cerca de la mitad de los átomos se desintegrarán dentro de los 3,82 días.

Las leyes de forma estadística básica pueden considerarse como equivalentes menos restrictivos de las leyes que tienen la forma condicional universal

$$(x)\ (Fx \supset Gx)$$

y afirman que todo caso de F es un caso de G, por ejemplo: "todo gas se expande cuando se lo calienta a presión constante". En realidad, los dos tipos de leyes comparten una característica importante, que es sistomática de su carácter nomológico: ambos hacen afirmaciones generales concernientes a una clase de casos que pueden considerarse como potencialmente infinitos. Como observamos antes, un enunciado lógicamente equivalente a una conjunción finita de proposiciones singulares y que, en este sentido, hace una afirmación concerniente sólo a una clase finita de casos no puede ser una ley y carece de la fuerza explicativa de un enunciado nomológico. Las oraciones legales, verdaderas o falsas, no son simplemente resúmenes convenientes de conjuntos finitos de datos relativos a casos particulares.

Por ejemplo, la ley de que los gases se dilatan al calentárselos a presión constante no equivale al enunciado de que, en todos los casos que han sido observados hasta ahora o quizás en todos los casos que se han producido hasta ahora, un aumento en la temperatura de un gas a presión constante ha ido acompañado por un aumento de volumen. Más bien afirma que el aumento de volumen está asociado al calentamiento de un gas a presión constante en *todos* los casos pasados, presentes o futuros, observados o no. Hasta implica condicionales contrafácticos y subjuntivos según los cuales si se hubiera calentado o se calentara una masa determinada de gas a presión constante, su volumen habría aumentado o aumentaría.

De manera similar, las leyes probabilísticas de la genética o de la

desintegración radiactiva no equivalen a informes descriptivos de las frecuencias con las cuales se ha hallado que ciertos tipos de fenómenos se producen en una clase finita de casos observados: afirman ciertos modos peculiares, es decir, probabilísticos, de conexión entre clases potencialmente infinitas de sucesos. En una ley estadística que especifica frecuencias relativas en un conjunto finito, no se supone finita la "clase de referencia" F. En realidad, podríamos decir que una ley de la forma "$p(G,F) = r$" no sólo se refiere a todos los casos reales de F sino también, por decirlo así, a la clase de todos sus casos potenciales. Supongamos, por ejemplo, que tenemos un tetraedro regular homogéneo cuyas caras están marcadas "I", "II", "III", "IV". Podemos afirmar, entonces, que la probabilidad de obtener un III haciendo rodar el tetraedro es de 1/4. Pero, si bien esta afirmación dice algo acerca de la frecuencia con la cual se obtiene un III al arrojar un tetraedro, no se la puede concebir simplemente como especificando esta frecuencia para la clase de todos los tiros que, de hecho, se hagan con el tetraedro. Pues podemos mantener nuestra hipótesis aunque se nos informe que el tetraedro fue arrojado realmente unas pocas veces en toda su existencia, y en este caso, sin duda, nuestro enunciado de probabilidad no se entendería en el sentido de afirmar que un cuarto, exacto o siquiera aproximado, de esos tiros deben dar el resultado III. Además, nuestro enunciado sería totalmente significativo y hasta podría hallarse bien sustentado (por ejemplo, por los resultados obtenidos con tetraedros similares o con otros cuerpos homogéneos con forma de sólidos regulares) aunque el tetraedro en cuestión fuera destruido sin habérselo arrojado nunca. Lo que el enunciado de probabilidad atribuye al tetraedro, pues, no es la frecuencia con la cual se obtiene el resultado III en lanzamientos reales pasados o futuros, sino una cierta *disposición*, a saber, la disposición a presentar el resultado III en aproximadamente uno de cada cuatro casos. Esta disposición puede caracterizarse por medio de una oración condicional subjuntiva: si se arroja el tetraedro un gran número de veces, se obtendría el resultado III en un cuarto de los casos, aproximadamente.[53] Así, las implicaciones en la forma de condicionales contrafácticos y subjuntivos son características de los enunciados legales tanto de forma estrictamente universal como estadística.

[53] Carnap (1951-54, págs. 190-192) ha sostenido, con espíritu similar, que la probabilidad estadística de sacar un as con un dado determinado es una característica física a la que también llama "el estado probabilístico" del dado, y que la frecuencia relativa con la cual el dado da un as es un síntoma de ese estado, así como la dilatación de la columna de mercurio de un termómetro es un síntoma de su estado térmico.

La concepción disposicional que he dado del concepto de probabilidad estadística parece también estar en íntimo acuerdo con la "interpretación por la propensión" propugnada por Popper. Esta última "difiere de la interpretación puramente estadística o frecuencial sólo en lo siguiente: que considera la probabilidad como una propiedad característica del ordenamiento experimental, no como la propiedad de una sucesión"; la propiedad aludida es explícitamente concebida como *disposicional* (Popper, 1957, págs. 67-68). Véase también la discusión de este artículo en Körner (1967, págs. 78-89, *passim*). Sin embargo, las enunciaciones corrientes de la interpretación basada en la propensión son todas bastantes breves; se hará una exposición más completa en un próximo libro de Popper.

En cuanto a la distinción entre oraciones legales de forma estrictamente universal y las de forma probabilística o estadística, se piensa a veces que los enunciados que afirman conexiones del primer caso, como la ley de Galileo o la ley newtoniana de la gravedad sólo se basan, al fin de cuentas, en un conjunto de elementos de juicio finito y, por ende, inevitablemente incompleto; que pueden tener, por lo tanto, excepciones aún no identificadas; y que, por consiguiente, también sólo deben calificarse como probabilísticos. Pero este argumento confunde la afirmación hecha por un enunciado dado con los elementos de juicio disponibles en su apoyo. A este último respecto, todos los enunciados empíricos se hallan sustentados en mayor o en menor grado por los elementos de juicio a nuestra disposición; o, según la terminología de algunos teóricos, estos elementos de juicio les confieren una probabilidad lógica o inductiva más o menos alta. Pero la distinción entre enunciados legales de forma estrictamente universal y los de forma probabilística alude no a los elementos de juicio que los sustentan, sino a las afirmaciones que hacen: hablando en términos aproximados, los primeros atribuyen (verdadera o falsamente) una característica a todos los miembros de una cierta clase; y los segundos, a una proporción específica de sus miembros.

Aun cuando todas las leyes supuestamente universales de la ciencia empírica lleguen a considerarse como reflejos de uniformidades estadísticas subyacentes —interpretación que la teoría cinética de la materia hace de las leyes clásicas de la termodinámica, por ejemplo— aun en tal caso, pues, no se borraría la distinción entre los dos tipos de leyes y las explicaciones correspondientes: de hecho, se la presupone en la formulación misma de la conjetura.

$$(x)\ (Fx \supset Gx)$$

Tampoco un enunciado de la forma condicional universal es lógicamente equivalente al enunciado correspondiente de la forma estadística básica

$$p(G, F) = 1$$

pues, como mostraremos con más detalle en la sección 3.3, el último sólo afirma que es prácticamente cierto que en un gran número de casos de F, casi todos son casos de G; por consiguiente, el enunciado probabilístico puede ser verdadero aunque el enunciado correspondiente de forma estrictamente universal sea falso.

Hasta ahora sólo nos hemos referido a leyes estadísticas de forma básica. Digamos ahora, con mayor generalidad, que *un enunciado tiene la forma de una ley estadística* o es de carácter probabilístico-estadístico si está formulado en términos de probabilidades estadísticas, es decir, si contiene (no vacuamente) el término "probabilidad estadística" o algún equivalente connotacional, o un término —como el de "vida media"— definido por medio de probabilidades estadísticas.

Tomemos como ejemplo el enunciado de que cuando se dejan caer simultáneamente dos monedas, la faz presentada por una de ellas es independiente de la presentada por la otra. Esto equivale a afirmar que la

probabilidad de que la segunda moneda salga cara cuando la primera sale cara es la misma que cuando la primera sale cruz, y viceversa. En general, las afirmaciones de independencia estadística tienen la forma de leyes estadísticas, si bien no son de forma estadística básica. Análogamente, un enunciado que afirme una dependencia o un "efecto posterior" estadístico tiene la forma de una ley estadística; por ejemplo, el enunciado de que, en una zona determinada, la probabilidad de que un día sea nublado cuando sigue a un día nublado es mayor que cuando sigue a un día despejado. Otras leyes de forma estadística, aun, se formulan en términos de valores medios de ciertas variables, como la energía cinética media y la trayectoria libre media de las moléculas de un gas; la noción de valor medio se define con referencia a probabilidades estadísticas.

Entendamos ahora por *explicación estadística* toda explicación que haga un uso esencial de al menos una ley o principio teórico de forma estadística. En las subsecciones siguientes, examinaremos la estructura lógica de tales explicaciones. Hallaremos que hay dos tipos lógicamente diferentes de explicación estadística. Uno de ellos consiste, básicamente, en la subsunción deductiva de una uniformidad estadística limitada bajo otras más amplias: lo llamaremos el tipo de *explicación deductivo-estadística.* El otro supone la subsunción, en un suceso particular bajo leyes estadísticas; por razones que daremos más adelante, lo llamaremos tipo de *explicación inductivo-estadística.*

3.2. *La explicación deductivo-estadística*

Un ejemplo de la llamada falacia del jugador es suponer que cuando varios tiros sucesivos de una moneda equilibrada han dado caras la probabilidad de que el tiro siguiente salga cruz es mayor que la probabilidad de que salga cara. Puede explicarse por qué esto no es así por medio de dos hipótesis que tienen la forma de leyes estadísticas. La primera es que el experimento de azar de arrojar una moneda equilibrada da caras con una probabilidad estadística de 1/2. La segunda hipótesis es que los resultados de diferentes lanzamientos de la moneda son estadísticamente independientes, de modo que la probabilidad de cualquier sucesión específica de resultados —tal como dos veces cara, luego cruz, luego cara, luego tres veces cara— es igual al producto de las probabilidades de los resultados aislados constituyentes. Estas dos hipótesis en función de probabilidades estadísticas implican *deductivamente* que la probabilidad de que salga cara después de una larga sucesión de caras es aún 1/2.

Ciertas explicaciones estadísticas de la ciencia tienen el mismo carácter deductivo, aunque a menudo son muy complejas matemáticamente. Consideremos, por ejemplo, la hipótesis de que los átomos de toda sustancia radiactiva tienen una probabilidad característica de desintegrarse durante un determinado intervalo de unidad de tiempo, y que esta probabilidad es independiente de la edad del átomo y de toda circunstancia externa. Esta hipótesis estadística compleja explica, por implicación deductiva, otros diversos aspectos estadísticos de la desintegración radiactiva, entre ellos, el siguiente: supongamos que se registra la desintegración de átomos individuales de una sustancia radiactiva por medio del centelleo

producido sobre una pantalla sensible por las partículas alfa emitidas por los átomos en desintegración. Entonces, los intervalos de tiempo que separan centelleos sucesivos variarán considerablemente en longitud, pero los intervalos de longitudes diferentes aparecerán con diferentes probabilidades estadísticas. De manera específica, si el intervalo de tiempo medio entre centelleos sucesivos es *s* segundos, entonces la probabilidad de que dos centelleos sucesivos estén separados por más de $n \cdot s$ segundos es $(1/e)^n$, donde *e* es la base de los logaritmos naturales.[54]

Las explicaciones de este tipo serán llamadas *deductivo-estadísticas* o *explicaciones D-E.* Suponen la deducción de un enunciado con la forma de una ley estadística a partir de un explanans que contiene indispensablemente por lo menos una ley o principio teórico de forma estadística. Se realiza la deducción por medio de la teoría matemática de la probabilidad estadística, que permite calcular ciertas probabilidades derivadas (las aludidas en el explanandum) sobre la base de otras probabilidades (especificadas en el explanans) halladas empíricamente o afirmadas hipotéticamente. Lo que explica una explicación D-E, pues, es siempre una uniformidad general expresada por una presente ley de forma estadística.

Por último, sin embargo, se pretende aplicar las leyes estadísticas a hechos particulares y establecer entre ellos conexiones explicativas y predictivas. En la subsección siguiente examinaremos la explicación estadística de sucesos particulares. Limitaremos nuestro examen al caso en el cual las leyes estadísticas explicativas son de forma básica: esto bastará para poner de manifiesto las diferencias lógicas básicas entre la explicación estadística y la deductivo-nomológica de hechos individuales.

3.3. *La explicación inductivo-estadística*

Como explicación de por qué el paciente Juan Pérez se recuperó de una infección por estreptococos se nos puede decir que a Pérez se le dio penicilina. Pero si tratamos de ampliar esta afirmación explicativa indicando una conexión general entre el tratamiento con penicilina y la remisión de una infección por estreptococos no podemos invocar de modo justificable una ley general según la cual la administración de penicilina produce la recuperación en todos los casos de tal infección. Lo que puede afirmarse, y lo que seguramente se da por supuesto aquí, es sólo que la penicilina logra la curación en un elevado porcentaje de casos, o con una alta probabilidad estadística. Este enunciado tiene el carácter general de una ley de forma estadística, y si bien no se especifica el valor de la probabilidad, el enunciado indica que ésta es alta. Pero, en contraste con los casos de explicaciones deductivo-nomológicas y deductivo-estadísticas, el explanans consiste en esta ley estadística junto con el enunciado de que el paciente recibió penicilina obviamente no implica el enunciado-explanandum, "el paciente se recuperó", con certeza deductiva, sino

[54] Véase Mises (1939, págs. 272-278). Allí se presentan tanto los hallazgos empíricos como el argumento explicativo. Este libro también contiene muchos otros ejemplos de lo que aquí llamamos explicación deductivo-estadística.

solamente, como suele decirse, con alta probabilidad o casi certeza. Dicho con mayor brevedad, pues, la explicación equivale a este razonamiento:

(3a) El caso particular de enfermedad de Juan Pérez, llamemos *j* a este caso, fue una grave infección por estreptococos (Ej) tratada con altas dosis de penicilina (Pj); y la probabilidad estadística $p(R, E \cdot P)$ de recuperación en los casos en los que E y P están presentes es cercana a 1; por consiguiente, era prácticamente seguro que se produciría la recuperación (Rj). Este razonamiento puede recibir la siguiente esquematización:

$p(R, E \cdot P)$ es cercana a 1

(3b) $Ej \cdot Pj$ (Por lo tanto:) Es prácticamente seguro (muy probable) que Rj

Así, en la literatura sobre la inferencia inductiva, los razonamientos basados en hipótesis estadísticas a menudo se han concebido como de esta forma u otra similar. Según esta concepción, es característico de la conclusión contener un operador modal tal como "casi ciertamente", "con alta probabilidad", "muy probablemente", etc. Pero la tesis de que los razonamientos tienen este carácter es insostenible. Pues las oraciones de la forma "es prácticamente cierto que *p*" o "es muy probable que *p*", donde el lugar de "*p*" está ocupado por algún enunciado, no son oraciones completas y autocontenidas que puedan calificarse como verdaderas o falsas. El enunciado que ocupa el lugar de "*p*" —por ejemplo, "*Rj*"— es verdadero o falso independientemente de todo elemento de juicio disponible, pero puede calificarse como más o menos probable, cierto, etc., sólo *con respecto a un conjunto de elementos de juicio.* El mismo enunciado, tal como "*Rj*", será cierto, muy probable, no muy probable, sumamente improbable, etc., según los elementos de juicio que se consideren. La frase "es casi seguro que *Rj*" en sí misma, pues, no es ni verdedera ni falsa; y no se la puede inferir de las premisas especificadas en (3b) ni de ningún otro enunciado.

La confusión subyacente en la esquematización (3b) puede ilustrarse también considerando su análoga para el caso de razonamientos deductivos. La fuerza de una inferencia deductiva, tal como la de "todos los *F* son *G*" y "*a* es *F*" para llegar a "*a* es *G*", se indica a veces diciendo que si las premisas son verdaderas, entonces la conclusión es necesariamente verdadera, formulación que podría sugerir la esquematización siguiente:

Todos los *F* son *G*; *a* es *F*
(Por lo tanto:) Es necesario (seguro) que *a* es *G*

Pero evidentemente las premisas dadas —que podrían ser, por ejemplo, "todos los hombres son mortales" y "Sócrates es hombre"— no establecen la oración "*a* es *G*" ("Sócrates es mortal") como una verdad necesaria o segura. La certeza a la que se alude en la paráfrasis informal del razonamiento es relacional: el enunciado "*a* es *G*" es seguro, o necesario, *con respecto a las premisas especificadas;* es decir, la verdad de éstas garantiza la verdad del primero, lo cual sólo significa que "*a* es *G*" es una consecuencia lógica de esas premisas.

Análogamente, presentar nuestra explicación estadística en la forma del esquema (3b) es concebir equivocadamente la función de las palabras "casi seguro" o "muy probable", tal como aparecen en la expresión formal de la explicación. Evidentemente, deben tomarse esas palabras en el sentido de que sobre la base de los elementos de juicio suministrados por el explanans, o con respecto a estos elementos de juicio, el explanandum es prácticamente seguro o muy probable, es decir, que:

(3c) "Rj" es prácticamente seguro (muy probable) con respecto al explanans que contiene las oraciones "$p(R, E \cdot P)$ es cercana a 1" y "$Ej \cdot Pj$".[55]

El razonamiento explicativo erróneamente representado por (3b), pues, podría esquematizarse de manera adecuada en la siguiente forma:

$$\text{(3d)} \quad \begin{array}{l} p(R, E \cdot P) \text{ es cercana a 1} \\ Ej \cdot Pj \\ \hline\hline Rj \end{array} \quad \text{[hace prácticamente seguro (muy probable)]}$$

En este esquema, la doble línea que separa las "premisas" de la "conclusión" significa que la relación de las primeras con la segunda no es de implicación deductiva, sino de apoyo inductivo, cuya fuerza está indicada en el corchete.[56, 57]

[55] Frases tales como "es casi seguro (muy probable) que *i* se recobrará", aun cuando se admita la concepción relacional aquí sugerida, manifiestamente conciernen a relaciones entre proposiciones, tales como las expresadas por las oraciones que forman la conclusión y las premisas de un razonamiento. Para los fines del presente examen, sin embargo, puede evitarse el compromiso con proposiciones concibiendo las frases aludidas como si expresaran relaciones lógicas entre *oraciones* correspondientes, por ejemplo, la oración-conclusión y la oración-premisa de un razonamiento. Esta concepción, que subyace en la formulación de (3c), será la adoptada en este ensayo, aunque por conveniencia ocasionalmente usaremos una paráfrasis.

[56] En la esquematización familiar de los razonamientos deductivos, donde una línea separa las premisas de la conclusión, no se hace ninguna distinción explícita entre una afirmación débil y otra fuerte, y puede entenderse cualquiera de ellas; o sea (*a*) que las premisas implican lógicamente las conclusión, y (*b*) que además las premisas son verdaderas. En el caso de nuestro razonamiento probabilístico, (3c) expresa una afirmación débil, análoga a (*a*); mientras que (3d) puede interpretarse como que expresa la "oferta de una explicación" (tomamos la expresión de Scheffler, 1957, sección 1) en la cual, además, las premisas explicativas se afirman —aunque tentativamente— como verdaderas.

[57] Las consideraciones esbozadas concernientes al uso de términos tales como "probablemente" y "ciertamente" en la forma de calificativos modales de enunciados individuales entran en conflicto también, según creo, con la noción de enunciado probabilístico categórico expuesta por C. I. Lewis en el siguiente pasaje (las bastardillas son del autor):

Así como "si *D* entonces (ciertamente) *P*, y *D* es el hecho" conduce a la consecuencia categórica "por lo tanto (ciertamente) *P*", así también "si *D* entonces probablemente *P*, y *D* es el hecho" conduce a una consecuencia categórica expresada por "es probable que *P*". Y esta conclusión no es meramente la enunciación, otra vez, de la relación de probabilidad entre "*P*" y "*D*", de igual modo que "por lo tanto (ciertamente) *P*" no es nuevamente la enunciación de "si *D* entonces (ciertamente *P*". "Si el barómetro está alto, mañana probablemente hará buen tiempo, y el barómetro está alto" afirma categóricamente algo expresado por "mañana proba-

Así, nuestra esquematización refleja explícitamente la comprensión de que "casi seguro", "muy probable", "prácticamente imposible" y otras expresiones similares usadas a menudo en la formulación de razonamientos probabilísticos, incluso explicaciones, no representan propiedades poseídas por ciertas oraciones o por las oraciones correspondientes, sino relaciones que ciertas oraciones tienen con otras. Según esta interpretación, la idea de que el explanans de (3d) hace al explanandum casi seguro o muy probable, sólo es un caso especial de la idea de que un enunciado o conjunto de enunciados dado —conjunto al que llamaremos los fundamentos o los elementos de juicio *e*— confiere apoyo, confirmación o credibilidad inductiva más o menos fuerte a un enunciado *h*. Por supuesto, aclarar y elaborar sistemáticamente la idea que aquí sólo caracterizamos de manera esquemática constituye el objetivo de diversas teorías del razonamiento inductivo. Aún está en discusión hasta qué punto es posible elaborar criterios claros y una teoría precisa del concepto en cuestión. Se han hecho diversos intentos para formular teorías lógicas rigurosas del concepto del apoyo inductivo que permita establecer graduaciones numéricas o no numéricas de su fuerza; dos ejemplos destacados de tales esfuerzos son la teoría de la probabilidad creada por Keynes y, especialmente, el impresionante sistema de lógica inductiva elaborada por Carnap.[58] En este último, el grado en el cual una oración, o hipótesis, *h*, está

blemente hará buen tiempo". Esta probabilidad es aún relativa a los fundamentos del juicio; pero si estos fundamentos son reales y contienen todos los elementos de juicio atinentes al caso disponibles, entonces no sólo es categórica sino que bien puede llamarse *la* probabilidad del suceso en cuestión (1946, pág. 319).

Esta posición me parece sujeta a las mismas objeciones indicadas en el texto. Si "*P*" es un enunciado, entonces las expresiones "ciertamente *P*" y "probablemente *P*", tales como se las considera en el pasaje citado, no son enunciados. Si nos preguntamos cómo haremos para discernir si son verdaderas, nos damos cuenta de que estamos en una total incertidumbre hasta que no se haya especificado un conjunto de enunciados o suposiciones y con respecto a los cuales pueda decirse que *P* es cierta, o altamente probable o ninguna de las dos cosas. Las expresiones aludidas, pues, son esencialmente incompletas; son formulaciones elípticas de enunciados relacionales y ninguna de ellas puede ser la conclusión de una inferencia. Por plausible que pueda parecer la sugerencia de Lewis, en la lógica inductiva no hay ningún análogo al *modus ponens*, o regla "de separación", de la lógica deductiva, según la cual dada la información de que "D" y "si *D* entonces *P*" son enunciados verdaderos, podemos separar el consecuente "*P*" de la premisa condicional y afirmarlo como enunciado aislado que debe ser también verdadero.

Al final del pasaje citado, Lewis sugiere la importante idea de que podría interpretarse "probablemente *P*" en el sentido de que los elementos de juicio totales atinentes al caso en el momento confieren a *P* una elevada probabilidad. Pero aun este enunciado es relacional, pues se refiere tácitamente a un tiempo no especificado y, además, Lewis no hace depender su noción de enunciado probabilístico categórico como conclusión de un razonamiento de la suposición de que las premisas del razonamiento incluyen todos los elementos de juicio disponibles atinentes al caso.

Debe señalarse, sin embargo, que en otras partes de su exposición Lewis destaca la relatividad de la probabilidad (lógica) y, por ende, da relieve a la misma característica que se opone a la idea de enunciados probabilísticos categóricos.

Objeciones similares, creo yo, se aplican a la concepción de Toulmin de los razonamientos probabilísticos; véase Toulmin (1958) y su discusión en Hempel (1960, secciones 1-3).

[58] Véase Keynes (1921). De los muchos escritos de Carnap sobre el tema, véase en particular (1954), (1950), 1952) y 1962).

confirmada por una proposición *e*, que es el elemento de juicio, está representado por una *c* (*h*, *e*), cuyos valores están comprendidos en el intervalo entre 0 y 1 inclusive, y que satisface todos los principios básicos de la teoría abstracta de la probabilidad; por ello, *c* (*h*, *e*) es llamada también la probabilidad *lógica* o *inductiva* de *h* sobre la base de *e*. Es menester distinguir claramente este concepto de la probabilidad inductiva como relación lógica cuantitativa entre enunciados, del concepto de probabilidad estadística como relación empírica cuantitativa entre tipos o clases de sucesos. Sin embargo, los dos conceptos tienen una estructura formal común, en virtud de la cual ambos pueden considerarse como probabilidades: ambos están definidos, en sus respectivas teorías formales, en términos de funciones aditivas no negativas cuyos valores están comprendidos entre 0 y 1. La teoría de Carnap brinda una definición explícita de *c* (*h*, *e*) para el caso en que las oraciones *h* y *e* pertenezcan a uno u otro de ciertos tipos relativamente simples de lenguajes formalizados; la extensión de este enfoque a lenguajes cuyo equipo lógico sea adecuado para la formulación de teorías científicas avanzadas es todavía un problema en discusión.

Pero independientemente de la medida en la cual la relación entre el explanandum y el explanans pueda analizarse en función del concepto cuantitativo de probabilidad inductiva elaborado por Carnap, las explicaciones probabilísticas deben considerarse como inductivas en el sentido amplio que hemos indicado. Para referirnos a la noción general de un apoyo inductivo que admita graduaciones sin comprometernos con ninguna teoría particular del apoyo o la configuración inductiva, usaremos la expresión "*(grado de) apoyo inductivo de h relativo a e*".[59]

Así, las explicaciones de hechos o sucesos particulares por medio de leyes estadístico-probabilísticas se presentan como razonamientos que son *inductivos* o *probabilísticos* en el sentido de que el explanans confiere al explanandum un grado más o menos alto de apoyo inductivo o de probabilidad lógica (inductiva); por ello, serán llamadas *explicaciones inductivo-estadísticas*, o *explicaciones I-E*. Las explicaciones en las cuales las leyes estadísticas invocadas son de forma básica, como (3d), serán llamadas también *explicaciones I-E de forma básica*.

Trataré de demostrar ahora que la concepción inductiva aquí sugerida de la explicación estadística de hechos particulares es también requerida por la interpretación empírica que han recibido las leyes probabilísticas en versiones de la teoría de la probabilidad estadística y sus aplicaciones.

La teoría matemática de la probabilidad estadística pretende suministrar una explicación teórica de los aspectos estadísticos de procesos repetitivos de cierto tipo, a los que se llama procesos de azar o experimen-

[59] Algunos intentos recientes para dar elucidaciones precisas de esta noción general han conducido a conceptos que no tienen todas las características formales de una función probabilística. Una de tales concepciones se encontrará en Helmer y Oppenheim (1945) y, menos técnicamente, en Hempel y Oppenheim (1945). Otra es el concepto de grado de apoyo fáctico, propuesta y desarrollada teóricamente por Kemeny y Oppenheim (1952). En Rescher (1958) se hallará una sugerente distinción y comparación de diferentes conceptos de elementos de juicio.

tos de azar. En términos aproximados, un experimento de azar es un tipo de proceso o suceso que puede ser repetido indefinidamente por el hombr o por la naturaleza y que, en cada caso, da un resultado de un cierto conjunto finito o infinito de "resultados", de manera que si bien los resultados varían de un caso a otro de un modo irregular e impredecible, las frecuencias relativas con las que aparecen los diferentes resultados tienden a ser más o menos constantes a medida que aumenta el número de hechos. El lanzamiento de una moneda, cuyos resultados posibles son cara o cruz, es un ejemplo familiar de un experimento de azar.

La teoría de la probabilidad ofrece un "modelo matemático" de las propiedades e interrelaciones matemáticas generales de las frecuencias a largo plazo asociadas con los resultados de los experimentos de azar.

En el modelo, cada uno de los diferentes "resultados posibles" asignados a un determinado experimento de azar F, está representado por un conjunto G, que puede concebirse como el conjunto de aquellas realizaciones del experimento que dan el resultado en cuestión, mientras que F puede considerarse como el conjunto de todas las realizaciones del experimento de azar. La probabilidad de obtener un tipo determinado G como resultado de realizar un experimento de tipo F se representa como una medida, $p_F(G)$, del tamaño del conjunto G en relación con el conjunto F.

Los postulados de la teoría matemática especifican que p_F es una función-conjunto aditiva y no negativa cuyo valor máximo es 1, es decir, para cada resultado posible G de F, $p_F(G) \geqslant 0$; si G_1, G_2 son resultados mutuamente excluyentes de F, entonces $p_F(G_1 \vee G_2) = p_F(G_1) + p_F(G_2)$; y $p_F(F) = 1$. Estas estipulaciones permiten probar los teoremas de la teoría elemental de la probabilidad; para tratar los experimentos que admiten infinitos resultados diferentes es adecuado extender el requisito de aditividad a sucesiones infinitas de conjuntos de resultados mutuamente excluyentes, $G_1, G_2, G_3, \ldots$

La teoría abstracta resultante puede aplicarse a cuestiones empíricas mediante una interpretación que relaciona enunciados en función de probabilidades como medidas de conjuntos con enunciados acerca de frecuencias relativas a largo plazo, asociadas con los resultados de experimentos de azar. Enunciaré ahora esta interpretación en una formulación que es, esencialmente, la dada por Cramér.[60] Por conveniencia, reemplazaremos en lo sucesivo la notación "$P_F(G)$" por "$p(G, F)$".

(3e) *Interpretación frecuencial de la probabilidad estadística.* Sea F un tipo determinado de experimento de azar y G un resultado posible del mismo; entonces, el enunciado $p(G, F) = r$ significa que en una larga serie de repeticiones de F, es prácticamente seguro que la frecuencia relativa del resultado G será aproximadamente igual a r.

Cramér también enuncia dos corolarios de esta interpretación que se refiere a aquellos casos en los que r difiere muy poco de 0 o de 1; son

[60] Véase Cramér (1946, págs. 148-149). El libro de Cramér contiene una detallada discusión de los fundamentos de la teoría de la probabilidad estadística y de sus aplicaciones. Formulaciones similares de la interpretación frecuencial han dado anteriores representantes de esta concepción, basada en la teoría de la medida, de la probabilidad estadística; por ejemplo, Kolmogoroff (1933, pág. 4).

de especial interés para nuestro examen posterior de la explicación probabilística. Por lo tanto, los destacaré aquí siguiendo nuevamente, en lo esencial, la formulación de Cramér.[61]

(3e.1) Si $1 - p(G, F) < \epsilon$, donde ϵ es un número positivo muy pequeño, entonces, si se realiza una sola vez el experimento de azar F, es prácticamente seguro que se producirá el resultado G.

(3e.2) Si $p(G,F) < \epsilon$, donde ϵ es un número positivo muy pequeño, entonces, si se realiza una sola vez el experimento de azar F, es prácticamente seguro que el resultado G no se producirá.

Como la interpretación frecuencial aquí formulada utiliza expresiones vagas como "una larga serie", "prácticamente seguro", "aproximadamente igual", etc., evidentemente no suministra una definición precisa de probabilidades estadísticas en función de frecuencias relativas observables. Pero parece inevitable cierta vaguedad para que el cálculo matemático de probabilidades pueda servir como representación teórica de las relaciones matemáticas entre frecuencias relativas establecidas empíricamente, que permanecen constantes sólo aproximadamente cuando aumenta la muestra observada.[62]

De particular interés para el análisis de la explicación *I-E*, sin embargo, es el hecho de que la frase "es prácticamente seguro que" aparezca en el enunciado general (3e) de la interpretación estadística y de que sus dos corolarios especiales (3e.1) y (3e.2) también contengan esta frase, aunque logren evitar las vagas expresiones "una larga serie de repeticiones" y "aproximadamente igual". La función de las palabras "es prácticamente seguro que" es clara: indican que la conexión lógica entre enunciados de probabilidad estadística y los enunciados de frecuencia empírica asociados a ellos es inductiva, no deductiva. Puede hacerse esto más explícito reformulando (3e) del siguiente modo: la información de que $p(G, F) = r$ y de

[61] Para (3e.1) véase Cramér (1946, pág. 150); para (3a.2) véase Cramér (1946, pág. 149), y una formulación muy similar en Kolmogoroff (1933, pág. 4).

[62] En ciertas formulaciones de la teoría matemática, la probabilidad estadística de un resultado determinado se define explícitamente como el límite de la frecuencia relativa de ese resultado en una serie infinita de realizaciones del experimento de azar corespondiente. Mises (1931 y 1939) y Reichenbach (1949) desarrollaron dos variantes importantes de este enfoque. Pero las series infinitas de ejecuciones no son realizables ni observables y, por ende, la definición basada en la idea de límite de la probabilidad estadística no suministra ningún criterio para la aplicación de este concepto a cuestiones empíricas observables. A este respecto, la concepción de la probabilidad basada en la idea de límite es un concepto teórico idealizado, y los criterios para su aplicación empírica tendrán que contener nuevamente algunos términos vagos del tipo que figura en (3c) y sus corolarios. En particular, un enunciado que especifique el límite de la frecuencia relativa del resultado G en una sucesión infinita de ejecuciones del experimento de azar F no tiene implicaciones deductivas concernientes a la frecuencia de G en un conjunto finito de ejecuciones, por grande que éste sea. La relación entre los enunciados probabilísticos así concebidos y los enunciados correspondientes acerca de las frecuencias relativas en conjuntos finitos debe, pues, ser considerada nuevamente como inductiva.

Una explicación concisa de la concepción basada en la idea de límite de la probabilidad estadística y un claro examen de algunas de sus dificultades se encontrará en Nagel (1939, en particular las secciones 4 y 7).

que C es un conjunto de n realizaciones de F, donde n es un número grande, confiere una casi certeza (alto apoyo inductivo) al enunciado de que el número de esas realizaciones de C cuyo resultado sea G es aproximadamente $n.r$. Los dos corolarios pueden concebir análogamente. Así, (3e.1) puede reformularse de la siguiente manera: la información de que $1 - p(G, F) < \epsilon$ (donde ϵ es un número positivo pequeño) y de que el hecho individual i es una realización del experimento de azar F (o de Fi, para abreviar) da fuerte apoyo inductivo al enunciado i que proporciona el resultado G, o que Gi, para abreviar. O bien, en una formulación un poco diferente: "Gi" es prácticamente seguro con respecto a las dos proposiciones "$p(G,F)$ es muy cercana a 1" y "Fi". Esta última versión tiene la misma forma que (3c); así, al dar una interpretación inductiva al contenido explicativo de las leyes probabilísticas de la manera que ilustra (3d), coincidimos básicamente con la interpretación empírica que da a las leyes probabilísticas la teoría contemporánea de la probabilidad estadística.[63]

En nuestro ejemplo concerniente a la recuperación de una infección por estreptococos, la ley estadística invocada no especificaba un valor numérico para la probabilidad de lograr la recuperación por medio de la penicilina. Ahora examinaremos un caso simple de explicación *I-E* en el cual el enunciado de probabilidad correspondiente es muy específico. Sea el experimento D (más exactamente, un experimento del *tipo* D) consistente en extraer una bolilla de un urna que contiene 999 bolillas blancas y una negra, todas del mismo tamaño y el mismo material, volviendo luego a colocar la bolilla en la urna. Podemos aceptar la hipótesis estadística de que, con respecto a los resultados "bolilla blanca" y "bolilla negra", D es un experimento de azar en el cual la probabilidad de obtener una bolilla blanca es $p(B, D) = 0{,}999$. De acuerdo con la interpretación estadística, es posible someter a prueba esta hipótesis mediante muestras estadísticas finitas, pero para nuestros propósitos actuales no necesitamos tomar en consideración los fundamentos que podamos tener para aceptar la hipótesis, pues sólo nos interesa aquí su uso explicativo. Nuestra regla (3e.1) sugiere que podría usarse la hipótesis, en verdad, para explicar probabilísticamente los resultados de ciertas extracciones individuales de la urna, es decir, los resultados de ciertas realizaciones de D. Supongamos, por ejemplo, que una extracción particular, d, da una bolilla blanca. Puesto que $p(B, D)$ difiere de 1 en menos de 0,0011, digamos, cantidad sumamente pequeña, la regla (3e.1) sugiere el siguiente razonamiento explicativo en analogía con (3d):

$$\text{(3f)} \qquad \frac{\begin{array}{c} 1 - p(B, D) < 0{,}0011 \\ Dd \end{array}}{Bd} \quad \text{[hace prácticamente seguro]}$$

[63] Sin embargo, los representantes de la actual teoría de la probabilidad estadística no toman en cuenta de manera explícita, en general, el carácter inductivo de su interpretación estadística de los enunciados de probabilidad. Menos aún tratan de analizar el concepto inductivo de certeza práctica, el cual obviamente queda fuera de la teoría matemática que constituye su principal interés.

Nuevamente, el explanans no implica lógicamente el explanandum; y el razonamiento no demuestra que admitiendo la verdad de los enunciados que figuran en el explanans, cabe esperar el fenómeno indicado en el explanandum "con seguridad". Más bien puede decirse que el razonamiento muestra que sobre la base de la información suministrada por el explanans, cabe esperar el suceso del explanandum con seguridad "práctica" o con muy alta probabilidad.

La concepción de Carnap de la lógica inductiva sugiere que la vaga expresión "hace prácticamente seguro", que aparece entre corchetes en (3f), podría reemplazarse por una expresión cuantitativa más definida. Esto requeriría la extensión de la teoría de Carnap a lenguajes en los que sea posible formular enunciados de probabilidad estadística. Aunque la lógica de los lenguajes tratados en la obra publicada de Carnap no es suficientemente rica para este propósito [64] parece indudable que en casos simples como los ilustrados en (3f), el valor numérico de la probabilidad lógica debe ser igual al de la correspondiente probabilidad estadística. Por ejemplo, la información de que la probabilidad estadística de extraer de la urna una bolilla blanca es de 0,999, y de que el hecho particular *d* es una extracción de la urna, debe conferir una probabilidad lógica de 0,999 a la "conclusión" de que la bolilla obtenida en *d* es blanca. Con mayor generalidad puede formularse esta regla de la manera siguiente:

(3g) si *e* es el enunciado "$(p(G,F) = r) \cdot Fb$", y *h* es "Gb", entonces $c(h,e) = r$.

Esta regla concuerda con la concepción, expuesta por Carnap, de la probabilidad lógica como un cociente equilibrado de apuestas para una apuesta a *h* sobre la base de *e*. Está de acuerdo igualmente con la concepción de Carnap de que la probabilidad lógica, sobre la base de los elementos de juicio *e*, de la hipótesis de que un caso particular *d* tendrá una propiedad especificada *M* puede ser considerada como una estimación, basada en *e*, de la frecuencia relativa de *M* en cualquier clase *K* de casos sobre los cuales no informan los elementos de juicio *e*. En realidad, Carnap agrega que la probabilidad lógica de "*Nb*" sobre la base de *e* en ciertos casos puede considerarse como una estimación de la probabilidad estadística de *m*.[65] Por lo tanto, si *e* realmente contiene la información de que la probabilidad estadística de *M* en *r*, entonces la estimación, sobre la base de *e*, de esta probabilidad estadística y, por ende, de la probabilidad lógica de "*Mb*" sobre la base de *e*, evidentemente debe ser también *r*.

Y así como la regla (3e.1) suministra la justificación lógica de las explicaciones estadísticas tales como (3f), así también nuestra regla (3g) suministra la justificación de un tipo similar de explicación probabilística, la cual invoca leyes estadísticas cuantitativamente definidas y que pueden esquematizarse del siguiente modo:

[64] Según una comunicación personal del profesor Carnap, éste ha extendido su sistema en tal dirección.

[65] Carnap (1950, págs. 168-175).

$$\text{(3h)}\quad \frac{\begin{array}{c} p(G,F)=r \\ Fi \end{array}}{Gi} \quad [r]$$

Un razonamiento explicativo de esta forma permitiría explicar el hecho de que un caso individual i presente la característica G, señalando que i es un caso de F; que la probabilidad estadística de que un F presenta la característica G es r; y que, de acuerdo con la regla (3g), esta información explicativa confiere la probabilidad lógica r al enunciado del explanandum. Me referiré a r también como a la probabilidad *asociada a* la explicación. Por supuesto, un razonamiento de este tipo sólo será considerado explicativo si el número r es bastante cercano a 1. Pero parece imposible designar, sin caer en la arbitrariedad, un número particular, por ejemplo, 0,8, como valor mínimo de la probabilidad r admisible en una explicación.

En nuestro ejemplo, puede darse a la explicación probabilística de la extracción de una bolilla blanca la forma (3h) del siguiente modo:

$$\text{3i)}\quad \frac{\begin{array}{c} p(B,D)=0{,}999 \\ Dd \end{array}}{Bd} \quad [0{,}999]$$

Ahora bien, se dice a menudo que las leyes probabilísticas pueden servir para explicar aspectos estadísticos de muestras extensas, pero no pueden explicar nada de un caso individual. Ejemplos como los siguientes parecen sustentar esta afirmación. La ley de que el lanzamiento de una moneda regular de caras con la probabilidad ½, evidentemente no nos permite explicar por qué un tiro particular sale cara; en cambio, la misma ley (más la suposición de que los resultados de diferentes tiros son estadísticamente independientes unos de otros) puede usarse para explicar el hecho de que el número de caras obtenidas en una serie de 10.000 tiros esté comprendido entre 4900 y 5100; pues este resultado tiene una probabilidad que pasa de 0,95. Pero si consideramos explicado este resultado debido a la alta probabilidad que el explanans le confiere, entonces, evidentemente, también debemos otorgar status explicativo a razonamientos como el de (3i), cuyo explanans hace altamente probable que el resultado dado se produzca si se realiza una sola vez el correspondiente experimento de azar.

A veces se piensa también que como los razonamientos probabilísticos no son lógicamente concluyentes, no pueden servir para explicar; pues aunque el explanans sea verdadero, es posible que el fenómeno del explanandum no se haya producido:[66] en el caso de (3i), por ejemplo,

[66] Así, Scriven (1959, pág. 467) dice que "los enunciados estadísticos son demasiado débiles, pues renuncian a dar cuenta del caso individual... Un suceso puede oscilar dentro de una red de leyes estadísticas". Dray (1963, pág. 119) expresa una opinión similar. Estas observaciones son muy correctas si se las entiende en el sen-

extraer *d* podría haber dado una bolilla negra a pesar de la alta probabilidad de extraer una bolilla blanca. Pero esta objeción a la idea de explicación probabilística reposa en una concepción demasiado limitada de la explicación científica, pues muchas explicaciones importantes de la ciencia empírica utilizan de modo totalmente explícito las leyes estadísticas que, junto con el resto de la información explicativa aducida, sólo hacen al explanandum altamente probable.

Por ejemplo, mediante los principios genéticos mendelianos puede demostrarse que es altamente probable que en una muestra al azar tomada de una población de plantas de guisantes, cuyas plantas progenitoras representan una cruza de una cepa pura de flores blancas y una cepa pura de flores rojas, aproximadamente el 75 % de las plantas tendrán flores rojas y las restantes, flores blancas. Este razonamiento, que puede utilizarse con propósitos explicativos o predictivos, es inductivo-estadístico; lo que explica o predice son los porcentajes aproximados de plantas de flores rojas y de plantas de flores blancas de la muestra. Las "premisas" con referencia a las cuales se demuestra que los porcentajes especificados son altamente probables incluyen: 1) las correspondientes leyes de la genética, algunas de las cuales tienen forma estadística y otras, estrictamente universal, y 2) información del tipo mencionado antes acerca de la constitución genética de la generación progenitora de las plantas de las que se toma la muestra. Los principios genéticos de forma estrictamente universal incluyen las leyes según las cuales los colores en cuestión están ligados a genes específicos, el gene rojo es dominante sobre el blanco y otras leyes generales diversas concernientes a la transmisión, por los genes, de los colores o, quizá, de un conjunto más amplio de caracteres vinculados genéticamente. Entre las generalizaciones estadísticas invocadas se encuentra la hipótesis de que las cuatro combinaciones posibles de genes determinantes del color —*BB*, *BR*, *RB*, *RR*— son estadísticamente equiprobables en su aparición en la progenie de dos plantas de la generación híbrida.

Examinemos ahora con un poco más de detenimiento el uso explicativo de la ley de la desintegración radiactiva del radón, según la cual este elemento tiene una vida media de 3,82 días. Puede invocarse esta ley para una explicación estadística del hecho de que a los 7,64 días una muestra particular consistente en 10 miligramos de radón se redujo, por desintegración radiactiva, a una cantidad residual que cae en algún punto del intervalo comprendido entre 2,4 y 2,6 miligramos; de manera similar, se la puede usar para predecir un resultado particular de este tipo. El quid del razonamiento explicativo y predictivo es el siguiente: el enunciado referente a la vida media del radón transmite dos leyes estadísticas: I) la probablidad estadística de que un átomo de radón sufra la desintegración radiactiva dentro de un período de 3,82 días es de ½, y II) las desintegraciones de átomos de radón diferentes constituyen hechos esta-

---

tido de que las leyes estadísticas no tienen implicaciones deductivas concernientes a sucesos particulares, pero son engañosas si se las usa para insinuar que las leyes estadísticas no pueden tener ninguna significación explicativa con respecto a sucesos particulares.

dísticamente independientes. Otra premisa utilizada es el enunciado según el cual el número de átomos que hay en 10 miligramos de radón es enormemente grande (pasa de 10^{19}). Como enseña la teoría matemática de la probabilidad, las dos leyes junto con este último enunciado implican deductivamente que es muy alta la probabilidad estadística de que la masa de átomos de radón sobrevivientes después de 7,64 días no difiere de 2,5 miligramos en más de 0,1 miligramos, es decir que caerá dentro del intervalo especificado. Más explícitamente, la consecuencia deducible de las dos leyes estadísticas junto con la información sobre el gran número de átomos implicados es otra ley estadística que dice: es muy alta la probabilidad estadística de que el experimento de azar *F* consistente en que 10 miligramos de radón se desintegren durante 7,68 días, dé un resultado de tipo *G*, a saber, una cantidad residual de radón cuya masa esté comprendida en el intervalo entre 2,4 y 2,6 miligramos. En verdad, la probabilidad es tan alta que, según la interpretación (9.2b), si se efectúa el experimento *F* una sola vez es "prácticamente seguro" que el resultado será de tipo *G*. En este sentido, es racional, sobre la base de la información dada, esperar que se produzca el resultado *G* en una sola realización en una sola realización de *F*. También en este sentido, la información concerniente a la vida media del radón y al gran número de átomos implicados en un experimento de tipo *F* suministra una explicación o predicción estadística de la producción de *G* en una realización particular del experimento.

Para dar otro ejemplo, tomemos el problema de explicar ciertos aspectos cuantitativos del movimiento browniano que presentan las pequeñas partículas suspendida en un líquido, fenómeno explicado cualitativamente como debido a los impactos irregulares, sobre las partículas en suspensión, de las moléculas circundantes en estado de agitación térmica. A partir de suposiciones basadas en los principios probabilísticos de la teoría cinética del calor, Einstein dedujo una ley según la cual el desplazamiento medio de tales partículas es proporcional a la raíz cuadrada del tiempo transcurrido.[67] Pero la definición teórica del desplazamiento medio está formulada en función de las probabilidades estadísticas de los diversos desplazamientos posibles, por lo cual la ley de Einstein tiene carácter probabilístico. Por consiguiente, no implica lógicamente valores definidos del desplazamiento medio que presentan números finitos de partículas. Pero la ley establece como altamente probable, en el sentido mencionado antes, que los desplazamientos medios de muestras finitas sean muy aproximadamente proporcionales a la raíz cuadrada del tiempo transcurrido, y se ha hallado que esto es lo que sucede de hecho. Así, la ley de Einstein brinda una explicación probabilística de ciertos aspectos observados del movimiento browniano.

Como ilustran estos ejemplos y otros que consideraremos pronto, las explicaciones en términos de leyes o teorías estadísticas cumplen un rol

[67] Mayores detalles y una descripción completa de algunos tests experimentales de esta fórmula se encontrarán en Svedberg (1912, págs. 89 y sigs.). Las ideas básicas de la explicación probabilística de otros aspectos cuantitativos del movimiento browniano se hallan lúcidamente expuestas en Mises (1939, pág. 259-268).

muy importante en la ciencia. En lugar de negarles status explicativo sobre la base de que la no realización del explanandum es compatible con el explanans, debemos reconocer que constituyen explicaciones de un carácter lógico diferente que reflejan, podríamos decir, un sentido diferente de la palabra "porque". Mises explica este punto de vista cuando anticipa, al contemplar cambios recientes en la noción de causalidad, que "la gente llegará gradualmente a sentirse satisfecha con los enunciados causales de este tipo: el 'seis' aparece con mayor frecuencia *porque* el dado está cargado (pero no sabemos cuál será el próximo número); o también: la radiación se hizo más intensa *porque* se aumentó el vacío y se incrementó el voltaje (pero no sabemos el número preciso de centelleos que se producirán en el minuto siguiente)".[68] Este pasaje se refiere claramente a la explicación estadística en el sentido que estamos considerando; expone lo que podría llamarse un concepto estadístico-probabilístico de "porque", en contraposición al concepto estrictamente determinante, que correspondería a la explicación deductivo-homológica.

Nuestro examen de la explicación estadística de sucesos particulares ha estado diirgido hasta ahora a poner de manifiesto su carácter inductivo. En la subsección siguiente, consideraremos otra importante característica que diferencia de modo tajante la explicación *I-E* de sus equivalentes deductivos.

3.4 *La ambigüedad de la explicación inductivo-estadística y el requisito de la máxima especificidad*

3.4.1. *El problema de la ambigüedad explicativa.* Consideremos una vez más la explicación (3d) de la recuperación en el caso particular *j* de la enfermedad de Juan Pérez. La ley estadística invocada en este caso afirma la recuperación como respuesta a la penicilina sólo para un alto porcentaje de infecciones con estreptococos, pero no para todas ellas; de hecho, ciertas cepas de estreptococos son resistentes a la penicilina. Digamos que un hecho, por ejemplo, un caso particular de enfermedad, tiene la propiedad S^* (o pertenece a la clase S^*) si es un caso de infección con una cepa de estreptococos resistente a la penicilina. Entonces, la probabilidad de recuperación entre casos elegidos al azar de S^* que sean tratados con penicilina será muy pequeña, es decir, $p(R, S^* \cdot P)$ será cercana a 0 y la probabilidad de la no recuperación, $p(R,S^* \cdot P)$ será cercana a 1. Pero supongamos ahora que la enfermedad de Pérez es, de hecho, una infección por estreptococos de la variedad resistente a la penicilina, y consideremos el siguiente razonamiento:

$$\text{(3k)} \quad \frac{\begin{array}{c} p(\overline{R},S^* \cdot P) \text{ es cercana a 1} \\ S^*j \cdot Pj \end{array}}{\overline{R}j} \quad \text{[hace prácticamente seguro]}$$

[68] Mises (1951, pág. 188). Bastardillas del autor.

Este razonamiento "rival" tiene la misma forma que (3d) y, según nuestras suposiciones, sus premisas son verdaderas, como lo son las de (3d). Sin embargo, su conclusión contradice la conclusión de (3d).

O bien supongamos que Pérez es un octogenario de corazón débil, y que en este grupo, S^{**}, la probabilidad de reponerse de una infección por estreptococos en respuesta a un tratamiento de penicilina, $p(R, S^{**} \cdot P)$, es muy pequeña. Entonces, tenemos el siguiente razonamiento rival de (3d), que presenta la no recuperación de Pérez como prácticamente segura a la luz de premisas que son verdaderas:

(3l)
$$\begin{array}{l} p(\bar{R}, S^{**} \cdot P) \text{ es cercana a } 1 \\ S^{**} j \cdot Pj \\ \hline\hline \bar{R}j \end{array} \quad \text{[hace prácticamente seguro]}$$

El peculiar fenómeno lógico ilustrado aquí será llamado la *ambigüedad de la explicación inductivo-estadística* o, más brevemente de la *explicación estadística.* Esta ambigüedad deriva del hecho de que un suceso individual determinado (por ejemplo, la enfermedad de Pérez) a menudo podrá obtenerse por selección al azar de una cualquiera entre varias "clases de referencia" tales como $S \cdot P, S^{*} \cdot P, S^{**}$
cuales el tipo de suceso (por ejemplo, R) ilustrado por el hecho dado tiene probabilidades estadísticas muy diferentes. Por consiguiente, para una explicación probalibística propuesta, con explanans verdadero que confiere la casi seguridad a un suceso particular, habrá a menudo un razonamiento rival de la misma forma probabilística y con premisas igualmente verdaderas que confiere la casi seguridad a la no producción del mismo hecho. Y toda explicación estadística de la producción de un suceso debe parecer sospechosa si existe la posibilidad de una explicación igualmente correcta, desde los puntos de vista lógico y empírico, de su no producción. *Esta categoría no tiene análoga en el caso de la explicación deductiva*; pues si las premisas de una explicación deductiva propuesta son verdaderas, entonces también es verdadera su conclusión; y su contradictoria, por ser falsa, no puede ser consecuencia lógica de un conjunto rival de premisas que sean igualmente verdaderas.

He aquí otro ejemplo de la ambigüedad de la explicación *I-E*: al expresar mi sorpresa por hallar que el tiempo, en Stanford, es cálido y soleado en un día de otoño, se me puede responder, a manera de explicación, que ello era de esperar porque la probabilidad de que el tiempo sea cálido y con sol (C) un día de otoño (N) en Stanford es, digamos, de 0.95. Esquemáticamente, esta explicación adoptaría la siguiente forma, donde 'n' representa "un día de otoño":

(3m)
$$\begin{array}{l} p(C, N) = 0{,}95 \\ Nn \\ \hline\hline Cn \end{array} \quad [0{,}95]$$

Pero supongamos que el día anterior fue frío y lluvioso, y que la probabilidad de que los sucesores inmediatos (S) de los días fríos y lluviosos, en Stanford, sean cálidos y soleados sea de 0,2. Entonces, la explicación (3m) tiene un rival en el siguiente razonamiento, el cual, con referencia a premisas igualmente verdaderas, presenta como casi seguro que este día de otoño no será cálido y soleado:

$$\text{(3n)} \qquad \begin{array}{c} p(\bar{C}, S) = 0{,}8 \\ Sn \\ \hline\hline \bar{C}n \end{array} \qquad [0{,}8]$$

De esta manera, el problema de la ambigüedad concierne a los razonamientos *I-E* cuyas premisas son verdaderas, no importa que seamos o no conscientes de ello. Pero, como indicaremos en seguida, el problema tiene una variante relacionada con las explicaciones cuyos enunciados-explanans, sean o no verdaderos de hecho, son *afirmados o aceptados* por la ciencia empírica en el momento en que se anuncia o se propone la explicación. Esta variante será llamada *el problema de la ambigüedad epistémica de la explicación estadística,* pues se relaciona con lo que se presume conocido en la ciencia, y no con lo que sucede de hecho, aunque sea desconocido para cualquiera.

Sea K_t la clase de todos los enunciados afirmados o aceptados por la ciencia empírica en el tiempo t. Esta clase representa, pues, la información científica total, o el "conocimiento científico" en el tiempo t. Usamos aquí la palabra "conocimiento" en el sentido con que hablamos comúnmente del conocimiento científico en un momento dado. No pretende afirmar que los elementos de K_t sean verdaderos ni que se sepa definidamente que lo sean. Ninguna afirmación semejante puede ser justificada a partir de ninguno de los enunciados establecidos por la ciencia empírica; y las normas básicas de la investigación científica exigen que un enunciado empírico, por bien sustentado que esté, sea aceptado y, por ende, admitido como miembro de K_t sólo a título de ensayo, es decir, sujeto a la condición de que se le pueda retirar tal privilegio si se descubrieran elementos de juicio desfavorables. Los miembros de K_t, pues, cambian en el curso del tiempo, pues como resultado de la investigación continua se admiten nuevos enunciados en la clase, mientras que otros pueden caer en el descrédito y ser abandonados. En lo sucesivo, llamaremos simplemente K a la clase de los enunciados aceptados cuando no sea necesaria ninguna referencia específica al tiempo en cuestión. Supondremos que K es lógicamente consistente y que constituye un conjunto cerrado con respecto a la implicación lógica, es decir que contiene todo enunciado implicado lógicamente por cualquiera de sus subconjuntos.

Podemos caracterizar ahora la *ambigüedad epistémica de la explicación I-E* del siguiente modo: el conjunto total de los enunciados científicos aceptados contiene diferentes subconjuntos de enunciados que pueden utilizarse como premisas en razonamientos de la forma probabilística que acabamos de considerar, y que confieren altas probabilidades a "conclu-

siones" lógicamente contradictorias. Nuestros ejemplos anteriores (3k), (3l), (3m) y (3n) ilustran este punto, si suponemos que las premisas de todos esos razonamientos pertenecen a *K*, en lugar de suponer que son todas verdaderas. Si se proponen uno o dos de tales razonamientos rivales con premisas en *K* como explicación de un hecho que la ciencia considera sucedido, entonces la conclusión del razonamiento, es decir, el enunciado-explanandum, por consiguiente, también pertenecerá a *K*. Y puesto que *K* es consistente la conclusión del razonamiento rival no pertenecerá a *K*. Sin embargo, es inquietante que podamos decir: independientemente de nuestra información acerca de si el suceso en cuestión (por ejemplo, tiempo cálido y soleado en Stanford, un día de otoño) se produjo o no, podemos dar una explicación del resultado informado en ambos casos; y además, una explicación cuyas premisas son enunciados científicamente establecidos que confieren una alta probabilidad lógica al resultado en cuestión.

Esta ambigüedad epistémica, digámoslo nuevamente, no tiene análogo en la explicación deductiva; puesto que *K* es lógicamente consistente, no puede contener conjuntos de premisas que impliquen conclusiones lógicamente contradictorias.

La ambigüedad epistémica también afecta al uso predictivo de los razonamientos estadísticos. En este caso, presenta el alarmante aspecto de ofrecernos dos razonamientos rivales con premisas científicamente bien establecidas, uno de los cuales caracteriza determinado suceso futuro como prácticamente seguro, mientras que el otro lo caracteriza como prácticamente imposible. ¿En cuál de esos razonamientos antagónicos podemos confiar racionalmente para la explicación o la predicción?

3.4.2. *El requisito de máxima especificación y la relatividad epistémica de la explicación inductivo-estadística.* Nuestras ilustraciones de la ambigüedad explicativa sugieren que la decisión sobre la aceptabilidad de una explicación o una predicción probabilística propuesta deberá tomarse a la luz de toda la información atinente a la cuestión que podamos obtener. Esto es lo que indica también un principio general cuya importancia para el razonamiento inductivo ha sido reconocido, aunque no siempre muy explícitamente, por muchos autores y que ha sido destacado por Carnap, quien lo llama *el requisito de los elementos de juicio totales.* Carnap lo formula del siguiente modo: "En la aplicación de la lógica inductiva a una situación cognoscitiva determinada, deben tomarse en consideración los elementos de juicio totales como base para determinar el grado de confirmación".[69] La utilización de sólo una parte de los ele-

[69] Carnap (1950, pág. 211). Se sugiere tal requisito, por ejemplo, en el pasaje de Lewis (1946) citado en la nota 57. Análogamente, Williams habla de "la más importante de todas las reglas de la lógica de la probabilidad, según la cual 'la' probabilidad de cualquier proposición es su probabilidad en relación con las premisas conocidas, y con ellas solamente" (Williams, 1947, pág. 72).

Agradezco mucho al profesor Carnap haberse señalado en 1945, cuando observé por primera vez la ambigüedad de los razonamientos probabilísticos, que ésta sólo es una de varias paradojas aparentes de la lógica inductiva que resultan de pasar por alto el requisito de los elementos de juicio totales.

Barker (1957, págs. 70-78) ha hecho una clara exposición original de la ambi-

mentos de juicio totales es permisible si el cotejo de los elementos de juicio carece de importancia para la "conclusión" inductiva, es decir, si sobre la base de los elementos de juicio parciales, la conclusión tiene la misma confirmación, o probabilidad lógica, que sobre la base de los elementos de juicio totales.[70]

El requisito de los elementos de juicio totales no es un postulado ni un teorema de la lógica inductiva; no se relaciona con la validez formal de los razonamientos inductivos. Más bien, como ha señalado Carnap, es una regla para la *aplicación* de la lógica inductiva; podríamos decir que enuncia una condición necesaria de la racionalidad de una aplicación semejante en una "situación cognoscitiva" dada, a la cual concebiremos como representada por el conjunto K de todos los enunciados aceptados en la situación.

Pero, ¿de qué manera debe relacionarse la idea básica de este requisito con la explicación probabilística? Sin duda, no debemos insistir en que el explanans contenga toda la información empírica disponible en ese momento y sólo ella. Ni tampoco *toda* la información disponible porque, de lo contrario, todas las explicaciones probabilísticas aceptables en el tiempo t deberían tener el mismo explanans, K_t; ni *solamente* la información disponible, porque una explicación propuesta puede satisfacer el requisito al no pasar por alto ninguna información disponible de importancia, no obstante lo cual puede invocar en el explanans algunos enunciados que aún no han sido sometidos a prueba lo suficiente como para ser incluidos en K_t.

Consideraciones como las que expondremos a continuación sugieren la medida en la cual debe imponerse el requisito de los elementos de juicio totales a las explicaciones estadísticas. Una explicación propuesta del restablecimiento de Pérez basada en la información de que Pérez tuvo una infección por estreptococos y fue tratado con penicilina, y de que la probabilidad estadística de recuperación en tales casos es muy alta, es inaceptable si incluye la información adicional de que los estreptococos de Pérez son resistentes a la penicilina o de que Pérez es un octogenario de corazón débil, y que e nestas clases de referencias la probabilidad de restablecimiento es pequeña. En realidad, sería deseable que una explicación aceptable se basara en un enunciado de probabilidad estadística perteneciente a la más restringida clase de referencia de la cual sea miembro el hecho particular en consideración, según nuestra información total. Así, si K nos dice no solamente que Pérez tuvo una infección por estreptococos y fue tratado con penicilina, sino también que es un octogenario de corazón débil (y si K no brinda ninguna información más específica que ésta), entonces exigiríamos que una explicación aceptable de la respuesta

güedad básica de los razonamientos probabilísticos, así como una apreciación escéptica acerca del requisito de los elementos de juicio totales como medio para abordar el problema. Sin embargo, en seguida propondré una manera de remediar la ambigüedad de la explicación probabilística con ayuda de una versión muy modificada del requisito de los elementos de juicio totales. Lo llamaré el requisito de la especificidad, y no será pasible de la misma crítica.

[70] Véase Carnap (1950, pág. 211 y pág. 494).

de Pérez al tratamiento se basara en una base estadística que formulase la probalilidad de esta respuesta en la más restringida clase de referencia a la cual nuestra información total asigna la enfermedad de Pérez, es decir, la clase de las infecciones por estreptococos que sufren los octogenarios de corazón débil.[71]

Permítaseme ampliar esta sugerencia remitiéndome a nuestro anterior ejemplo concerniente al uso de la ley según la cual la vida media del radón es de 3,82 días para explicar el hecho de que la cantidad residual de radón a la que queda reducida una muestra de 10 miligramos en 7,64 días está comprendida entre 2,4 y 2,6 miligramos. Según el conocimiento científico actual, la tasa de desintegración de un elemento radiactivo sólo depende de su estructura atómica, caracterizada por su número atómico y su número de masa, es decir, que no la afectan la edad de la muestra ni factores tales como la temperatura, la presión, las fuerzas magnéticas y eléctricas y las interacciones químicas. Así, al especificar la vida media del radón, la masa inicial de la muestra y el intervalo de tiempo en cuestión, el explanans toma en consideración toda la información disponible que es importante para estimar la probabilidad del resultado mediante leyes estadísticas. Podemos expresar esto de una manera un poco diferente: en las circunstancias supuestas, nuestra información total K asigna el caso en estudio, ante todo, a la clase de referencia F_1, digamos, de los casos en los que se deja desintegrarse durante 7,64 días una muestra de radón de 10 miligramos; y la ley de la vida media del radón asigna una probabilidad muy alta, dentro de F_1, al "resultado", digamos G, consistente en el hecho de que la masa residual de radón esté comprendida entre los 2,4 y los 2,6 miligramos. Supongamos ahora que K también contiene información acerca de la temperatura de la muestra dada, acerca de la presión y la humedad relativa bajo las cuales se la conserva, acerca de las condiciones eléctricas y magnéticas circundantes, etc., de modo que K asigna el caso dado a una clase de referencia mucho más restringida que F_1, digamos, $F_1\ F_2\ F_3\ \ldots\ F_n$. Ahora bien, la teoría de la desintegración radiactiva, también incluida en K, nos dice que la probabilidad estadística de G dentro de esa clase más estrecha es la misma que dentro de F_1. Por esta razón, en nuestra explicación basta con admitir la probabilidad $p(G, F_1)$.

Observemos, sin embargo, que son concebibles "situaciones cognoscitivas" en las cuales el mismo razonamiento no sería una explicación aceptable; supongamos, por ejemplo, que en el caso de la muestra de ra-

[71] Esta idea está íntimamente relacionada con la usada por Reichenbach (véase 1949, sección 72) en el intento por demostrar que es posible asignar probabilidades a sucesos particulares dentro del marco de una concepción estrictamente estadística de la probabilidad. Reichenbach proponía que la probabilidad de un suceso aislado, tal como la realización exitosa de un vuelo programado de determinado avión comercial, fuera concebida como la probabilidad estadística que el *tipo* de suceso considerado (la realización exitosa de un vuelo) posee dentro de la clase de referencia más reducida a la que pertenece el caso dado (el vuelo especificado del avión dado) y de la cual se dispone de una información estadística confiable (por ejemplo, la clase de los vuelos programados realizados hasta ahora por los aviones de la línea a que pertenece el avión dado, y en condiciones climáticas similares a las prevalecientes en el momento del vuelo en cuestión).

dón en estudio se midiera la cantidad restante una hora antes del final del período de 7,64 días y se hallara que es de 2,7 miligramos, con lo cual excedería bastante los 2,6 miligramos, hecho que, considerando la ley sobre la desintegración del radón, es altamente improbable pero no imposible. Este hallazgo, que luego formaría parte de los elementos de juicio totales *K*, asigna el caso particular en cuestión a una clase de referencia, digamos F^*, dentro de la cual, según la ley de desintegración del radón, el resultado *G* es altamente improbable, ya que sería necesario un ritmo desusado en la desintegración de la muestra dada para reducir los 2,7 miligramos de la última hora de la prueba a una cantidad comprendida entre 2,4 y 2,6 miligramos. Por consiguiente, la información adicional aquí considerada no puede despreciarse, y una explicación del resultado observado sólo será aceptable si tiene en cuenta la probabilidad de *G* en la clase de referencia más restringida, es decir, $p(G, F_1, F^*)$. (La teoría de la desintegración radiactiva implica que esta probabilidad es igual a $p(G, F^*)$, y como consecuencia de esto no es necesario tener en cuenta explícitamente la pertenencia del caso dado a F_1.)

Ahora podemos enunciar más explícitamente el requisito sugerido por las consideraciones precedentes; lo llamaremos el *requisito de la máxima especificidad para las explicaciones inductivo-estadísticas*. Consideremos una explicación de forma estadística básica:

$$\text{(3o)}\qquad \frac{\begin{array}{c} p(G, F) = r \\ Fb \end{array}}{Gb} \; [r]$$

Sea *s* la conjunción de las premisas, y si *K* es el conjunto de todos los enunciados aceptados en el tiempo dado, sea *k* una proposición lógicamente equivalente a *K* (en el sentido de que k es implicada por *K* y, a su vez, implica toda oración de *K*). Entonces, para ser racionalmente aceptable en la situación cognoscitiva representada por *K*, la explicación propuesta (3o) debe satisfacer la siguiente condición (el requisito de máxima especifidad): si $s \cdot k$ implica[72] que *b* pertenece a una clase F_1, entonces $s \cdot k$ debe también implicar un enunciado que especifique la probabilidad estadística de *G* en F_1, digamos:

$$p(G, F_1) = r_1$$

Aquí, r_1 debe ser igual a *r*, a menos que el enunciado de probabilidad citado sea simplemente un teorema de la teoría matemática de la probabilidad.

72 Es necesaria la referencia a $s \cdot k$ y no a *k* porque, como observamos antes, no concebimos la condición que estamos examinando como si requiriera que todos los enunciados-explanans aducidos sean aceptados por la ciencia en el momento en cuestión y, por ende, que estén incluidos en la correspondiente clase *K*.

La cláusula "a menos que" añadida es muy adecuada, y su omisión tendría consecuencias indeseables. Es muy adecuada porque los teoremas de la teoría matemática pura de la probabilidad no pueden brindar una explicación de cuestiones empíricas. Por lo tanto, se los puede descartar cuando investigamos si $s \cdot k$ puede darnos leyes estadísticas que especifiquen la probabilidad de G en clases de referencia más restringidas que F. Y la omisión de la cláusula provocaría inconvenientes, pues si se propone (3o) como explicación, entonces presumiblemente se acepta como un hecho que Gb; por consiguiente, "Gb" pertenece a K. Así, K asigna b a la clase más restringida $F \cdot G$, y en lo concerniente a la probabilidad de G en esta clase, s · k implica trivialmente el enunciado según el cual $p(G, F \cdot G) = 1$, que es sólo una consecuencia de los postulados teóricos de la medición de la probabilidad estadística. Puesto que $s \cdot k$ implica un enunciado de probabilidad más específico para G que el invocado en (3o), éste violaría el requisito de la máxima especificidad —como lo violaría toda explicación estadística propuesta para un suceso que consideramos ocurrido— de no ser por la cláusula "a menos que", la cual, en efecto, desautoriza la idea de que el enunciado "$p(G, F \cdot G) = 1$" suministra una ley más adecuada para explicar el presunto hecho de Gb.

El requisito de la máxima especificidad, pues, es presentado aquí a título de ensayo para caracterizar la medida en la cual el requisito de los elementos de juicio totales se aplican adecuadamente a las explicaciones inductivo-estadísticas. La idea general sugerida equivale a lo siguiente: al formular o juzgar una explicación I-E debemos tener en cuenta toda la información suministrada por K que tenga potencial importancia *explicativa* para el hecho del explanandum; es decir, todas las leyes estadísticas que pueden vincularse mediante las leyes estadísticas con el hecho del explanandum.[73]

El requisito de la máxima especificidad elimina el problema de la ambigüedad epistémica; pues puede verse fácilmente que de dos razonamientos estadísticos rivales que tienen asociadas altas probabilidades y premisas pertenecientes todas a K, uno de ellos al menos viola el requisito de la máxima especificidad. En efecto, sean

[73] Por basarse en esta idea general, y específicamente en el requisito de la máxima especificidad, el método aquí sugerido para eliminar la ambigüedad epistémica de la explicación estadística difiere sustancialmente del enfoque con el que en un estudio anterior (Hempel, 1962, en particular sección 10) abordé el mismo problema. En ese estudio, en el cual no distinguía explícitamente los dos tipos de ambigüedad explicativa caracterizados antes en esta sección, apliqué el requisito de los elementos de juicio totales a las explicaciones estadísticas de una manera que presuponía que el explanans de toda explicación aceptable pertenecía a la clase K, y exigía por tanto que la probabilidad que el explanans asigna al explanandum fuera igual a la que otorgan a éste los elementos de juicio totales, K. Los argumentos expuestos en esta sección indican las razones por las cuales este enfoque me parece insatisfactorio. Obsérvese, en particular, que si se lo aplica estrictamente, el requisito de los elementos de juicio totales excluye la posibilidad de obtener explicaciones estadísticas significativas de sucesos cuya realidad se considera un hecho establecido en la ciencia; pues toda oración que describa tal suceso es por lógica implicada por K y, por consiguiente, tiene trivialmente la probabilidad lógica 1 con respecto a K.

$$\frac{\begin{array}{c} p(G,F) = r_1 \\ Fb \end{array}}{Gb} \; [r_1] \qquad \text{y} \qquad \frac{\begin{array}{c} p(\overline{G},H) = r_2 \\ Hb \end{array}}{\overline{G}b} \; [r_2]$$

los razonamientos en cuestión, donde r_1 y r_2 son cercanas a 1. Entonces, puesto que *K* contiene las premisas de ambos razonamientos, asigna *b* a *F* y a *H*, y también, por ende, a $F \cdot H$. Por consiguiente, si ambos razonamientos satisfacen el requisito de la máxima especificidad, *K* debe implicar

$$p(G, F \cdot H) = p(G, F) = r_1$$
$$p(\overline{G}, F \cdot H) = p(\overline{G}, H) = r_2$$

pero $\quad p(G, F \cdot H) + p(\overline{G}, F \cdot H) = 1$

por consiguiente $\quad r_1 + r_2 = 1$

que es una falsedad aritmética, ya que r_1 y r_2 son ambos cercanos a 1; luego, no puede ser implicada por la clase consistente *K*.

Así, el problema de la ambigüedad epistémica ya no se plantea para las explicaciones I-E que satisfacen el requisito de la máxima especificidad. *Nunca* estamos en condiciones de afirmar: haya sucedido o no este hecho particular podemos dar una explicación aceptable de su resultado; y una explicación, además, cuyas premisas son enunciados aceptados científicamente que confieren una alta probabilidad lógica al resultado en cuestión.

Pero si bien se ha resulto el problema de la ambigüedad epistémica, nuestro requisito no afecta a la ambigüedad en el primer sentido examinado en esta sección; es decir, queda en pie el hecho de que dado un razonamiento estadístico con premisas verdaderas y una alta probabilidad asociada a él, puede existir un razonamiento rival con premisas igualmente verdaderas y que tenga también una alta probabilidad, cuya conclusión contradice la del primer razonamiento. Y aunque el conjunto *K* de enunciados aceptados en un momento cualquiera nunca incluye todos los enunciados que, de hecho, son verdaderos (e incluye, sin duda, muchos que son falsos), es perfectamente posible que *K* contenga las premisas de dos de tales razonamientos antagónicos; pero, como hemos visto, al menos uno de éstos no será aceptable porque viola el requisito de especificidad.

Las consideraciones precedentes demuestran que *el concepto de explicación estadística de sucesos particulares está esencialmente relacionado con una situación cognoscitiva dada, representada por una clase K de enunciados aceptados.* En verdad, el requisito de la máxima especificidad hace una referencia explícita e inevitable a esa clase, por lo cual sirve para caracterizar el concepto de "explicación I-E referida a la situación cognoscitiva representada por *K*". Llamaremos a esta característica la *relatividad epistémica de la explicación estadística.*

Podría parecer que el concepto de explicación deductiva posee el mismo tipo de relatividad, ya que la aceptabilidad de una explicación D-N o D-E dependerá no sólo de que sea deductivamente válida y haga un uso esencial del tipo adecuado de ley general, sino también de que sus premisas estén bien sustentadas por los elementos de juicio atinentes disponibles. Efectivamente, es así; y esta condición de confirmación empírica se aplica

igualmente a las explicaciones estadísticas que sean aceptables en una situación cognoscitiva dada. Pero la relatividad epistémica que el requisito de la máxima especificidad implica para las explicaciones I-E es de un tipo muy diferente y no tiene análoga en lo concerniente a las explicaciones D-N. Pues el requisito de especificidad no se refiere al apoyo que los elementos del juicio totales, *K*, suministran a los enunciados del explanans: no exige que éstos sean incluidos en *K*, ni siquiera que *K* suministre elementos de juicio en su apoyo. Más bien, se refiere a lo que podría llamarse el concepto de explicación estadística *potencial*. Pues estipula que, independientemente de los elementos de juicio que pueda haber en apoyo del explanans, una explicación I-E propuesta no es aceptable si su fuerza explicativa potencial con respecto al explanandum especificado está viciada por leyes estadísticas incluidas en *K* pero no en el explanans, y que permiten elaborar razonamientos estadísticos rivales. Como hemos visto, este peligro nunca surge en las explicaciones deductivas. Por consiguiente, éstas no están sujetas a ninguna condición restrictiva, y la noción de explicación deductiva potencial (en contraposición con una explicación deductiva de explanans bien confirmados) no requieren ninguna relativización con respecto a *K*.

Como consecuencia, podemos hablar con sentido de explicaciones D-N y D-E verdaderas: son aquellas explicaciones D-N y D-E cuyas premisas son verdaderas (y, por ende, también lo son sus conclusiones) independientemente de que esto sea sabido o creído y, por consigiuente, de que las premisas estén o no incluidas en *K*. Pero esta idea no tiene ningún análogo significativo en las explicaciones I-E, puesto que, como hemos visto, el concepto de explicación estadística exige la relativización con respecto a *K*.

3.4.3. *Sistemas de estados discretos y ambigüedad explicativa.* En un artículo claro e instructivo, Rescher[74] ha mostrado que ciertos sistemas físicos de un tipo particular, a los que llama sistemas de estados discretos, suministran excelentes ilustraciones de explicaciones y predicciones deductivas y probabilísticas, y que un examen más minucioso de dichos sistemas puede contribuir mucho a aclarar la estructura lógica, el alcance y las interrelaciones de tales procedimientos. Me propongo mostrar que un estudio de estos sistemas también nos enfrenta con el problema de la ambigüedad explicativa y da apoyo a la solución aquí sugerida.

Por *sistema de estados discretos*, o sistema E-D para abreviar, Rescher entiende un sistema físico que en cualquier momento dado está en uno de varios estados posibles, E_1, E_2, . . ., cada una de cuyas apariciones ocupa un período finito, aunque quizá muy breve; para nuestros propósitos, consideraremos que el número de estados posibles de un sistema E-D es finito. La sucesión de estados de un sistema E-D está gobernada por un conjunto de leyes, cada una de las cuales puede ser determinista o probabilística (estadística). Una ley determinista tiene la forma "el estado E_i es siempre seguido inmediatamente por el estado E_j"; una ley probabilística tiene la forma "la probabilidad estadística de que (una aparición de) el estado E_i sea seguido inmediatamente por (una aparición de) el

[74] Rescher (1963).

estado E_j es r_{ij}". Un sistema E-D de este tipo puede ser caracterizado por medio de la matriz de todas las probabilidades de transición r_{ij}.

Hay varios ejemplos físicos de sistemas E-D; entre ellos, Rescher menciona una computadora electrónica digital, un átomo de un elemento radiactivo en sus sucesivos estados de desintegración y —dado un modo de descripción adecuadamente esquematizado— una partícula en movimiento browniano. Una pelota que rueda por un tablero de Galton [75] es otro sistema E-D; su estado se encuentra representado, en un momento determinado, por la cantidad de alfileres que la separan horizontalmente de la línea vertical del centro del tablero.
momentáneo de un sistema E-D) como un razonamiento cuya conclusión

Rescher define una explicación probabilística potencial (de un estado tiene la forma "el estado del sistema en el intervalo de tiempo t es E_i" o "$est(t) = E_i$" para abreviar, y cuyas premisas consisten en las leyes que gobiernan el sistema y en un conjunto de enunciados que especifican los estados que presenta el sistema durante otros intervalos de tiempo, t_1, t_2, ..., t_n, todos los cuales son diferentes de t.[76] El razonamiento puede ser "probabilístico, o bien en el sentido fuerte ... en el cual $est(t) = E_i$ es (condicionalmente) *más probable que improbable*, o bien en el sentido débil ... en el cual $est(t) = E_i$ es (condicionalmente) *más probable que* $est(t) = E_j$ para todo $j \neq i$".[77] Por último, "un razonamiento potencialmente explicativo se convierte en una *explicación* (*real*) si sus premisas son reales o probablemente verdaderas".[78]

Para ver que la explicación probabilística así concebida está viciada nuevamente de ambigüedad, consideremos un sistema E-D de tres estados exactamente, E_1, E_2 y E_3, con probabilidades de transición como las que especifica el siguiente esquema:

	E_1	E_2	E_3
E_1	0	0,99	0,01
E_2	0	0	1
E_3	1	0	0

[75] Para un examen de este proceso, véase Mises (1939, págs. 237-240).

[76] Rescher no exige de una explicación potencial —como lo exige de una predicción potencial, que en otros aspectos se caracteriza de la misma manera— que los intervalos t_1, t_2, ..., t_n precedan todos a t. Como resultado de esto, toda predicción potencial es una explicación potencial, pero no a la inversa. En la sección 3.5 examinaremos las razones que aduce en apoyo de esta concepción.

[77] Rescher (1963, pág. 330), bastardillas del autor. El concepto de probabilidad condicional aquí introducido no recibe ulterior clarificación, pero evidentemente pretende representar la probabilidad que la conclusión de un razonamiento explicativo posee con respecto a las premisas o condicionadas por éstas. En este caso, las probabilidades condicionales tendrían el carácter general de las probabilidades lógicas; y Rescher parece operar con ellas de acuerdo con la concepción reflejada por nuestro esquema (3h), donde la "probabilidad" en cuestión se especifica entre corchetes que siguen a la doble línea que separa la conclusión de las premisas.

[78] Rescher (1963, pág. 329); bastardillas del autor.

Así, la probabilidad de que a E_1 le siga inmediatamente E_1 es 0; la de que le siga E_2, es 0,99; la de que le siga E_3 es 0,01; etcétera.

Alternativamente, los sistemas E-D pueden ser caracterizados por lo que Rescher llama diagramas de transición. En nuestro caso el diagrama adopta la siguiente forma:

E_1 (0,99) → E_2(1) → E_3(1) → E_1

E_1 (0,01) → E_3

Como puede verse fácilmente, las leyes de transición implican las dos leyes derivadas siguientes:

(L_1) La probabilidad de que el sucesor de dos períodos de E_1 sea E_3 es $0{,}99 \times 1 = 0{,}99$.

(L_2) La probabilidad de que el sucesor inmediato de E_3 sea nuevamente E_3 es 0.

Supongamos ahora que en dos intervalos de tiempo particulares sucesivos, t_1 y t_2, nuestro sistema presenta los estados E_1 y E_3, respectivamente; es decir, que los siguientes enunciados son verdaderos:

(C_1) $est(t_1) = E_1$

(C_2) $est(t_2) = E_3$

Entonces, C_1 junto con L_1 suministra las premisas para un razonamiento probabilístico que asigna la "probabilidad" 0,99 a la conclusión de que en el intervalo de tiempo t_3 que sigue inmediatamente a t_2, el sistema se encuentra en el estado E_3; es decir, que $est(t_3) = E_3$. Pero C_2 junto con L_2 da análogamente la probabilidad 1 a la conclusión de que $est(t_3) \neq E_3$. Según nuestras suposiciones, las premisas invocadas en estos razonamientos antagónicos son verdaderas; por consiguiente, los razonamientos constituyen explicaciones probabilísticas fuertes, en el sentido que les da Rescher, de la aparición y la no aparición de E_3 durante el intervalo de tiempo t_3; y ambos son explicaciones reales en el sentido que les otorga Rescher, ya que todas las premisas explicativas son verdaderas. Así, llegamos a la ambigüedad explicativa en el primero de nuestros dos sentidos. El hecho de que también hay ambigüedad en el segundo sentido, el epistémico, está claro si consideramos que según nuestras suposiciones todas las premisas invocadas pueden pertenecer, por supuesto, a la clase K de enunciados aceptados en el momento.[79]

Para excluir esta consecuencia inadmisible es necesario complementar las definiciones de Rescher de explicación y predicción probabilísticas con un adecuado requisito adicional. En nuestro ejemplo, el primero de los dos razonamientos rivales sería rechazado, evidentemente, sobre la base de que deja de lado informaciones importantes. Pero éste es, precisamente, el

[79] La misma ambigüedad dificulta el uso predictivo de estos razonamientos: aunque ambos se basan en premisas aceptadas (y verdaderas), conducen a predicciones contradictorias acerca del estado del sistema durante t_3.

veredicto del requisito de la máxima especificidad. Pues, en nuestra ilustración podemos suponer que la clase K incluye la información transmitida por C_1, C_2, L_1 y L_2; pero que K no contiene otra información específica que implique una asignación de probabilidad, sobre bases empíricas, a la proposición "$est(t_3) = E_3$". El primero de estos dos razonamientos probabilísticos viola el requisito de la máxima especificidad, ya que sólo tiene en cuenta que el estado del sistema en t_1 es E_1, aunque K nos dice además que a la aparición de E_1 en t_1 le sigue directamente una aparición de E_3, y que para una aparición de E_1 seguida por una aparición de E_3 la probabilidad de que E_3 sea el sucesor de dos períodos es 0. (Pues L_2 nos dice con toda generalidad que la probabilidad de que a una aparición de E_3 —cualquiera que sea su predecesor—le siga una aparición de E_3 es 0.) Por consiguiente, sólo el segundo de los dos razonamientos rivales es aceptable bajo el requisito de la máxima especificidad.

3.5. *Aspectos predictivos de la explicación estadística*

¿Puede sostenerse que una explicación inductivo-estadística de un hecho particular constituye, como la explicación deductivo-nomológica, una predicción potencial de ese suceso?

Si se incluye el enunciado que describe la aparición en cuestión en la clase K de los enunciados aceptados, entonces la cuestión de predecir el hecho evidentemente no puede plantearse en la situación cognoscitiva representada por K. Por ello, demos a nuestro problema esta forma: supongamos que un razonamiento del tipo (3o) satisface el requisito de la máxima especificidad relativa a K y que su explanans está bien confirmado por K; ¿sería aceptable, entonces, como razonamiento predictivo en la situación cognoscitiva caracterizada por K? La respuesta dependerá, por supuesto, de las condiciones que consideremos que debe satisfacer un razonamiento estadístico para ser racionalmente aceptable con propósitos predictivos en una situación cognoscitiva determinada. Examinemos brevemente esta cuestión.

La racionalidad exige, evidentemente, que al crearnos expectativas concernientes a los hechos futuros tomemos en consideración toda la información disponible en el momento: éste es el quid del requisito de los elementos de juicio reales. Pero, ¿cómo debe concebirse, más específicamente, este requisito? Si se dispone de una definición y una teoría generales de la probabilidad lógica, o inductiva, la condición equivale a lo siguiente: la probabilidad conferida a la conclusión del razonamiento predictivo por las premisas solamente debe ser igual a la probabilidad que le asignan los elementos de juicio totales K. En este caso, se justifica dejar de lado, en el razonamiento, la estimación de los elementos de juicio totales, pues su adición a las premisas no modificaría la probabilidad de la conclusión. En la actualidad, no se dispone de ninguna definición ni de ninguna teoría de la probabilidad inductiva suficientemente amplias como para ser aplicables a todos los tipos de razonamientos inductivos que sería menester considerar. Si se construyera tal definición, por ejemplo, generalizando

el enfoque de Carnap, podría resultar que un razonamiento estadístico cuyas permisas están bien sustentadas por K y que satisfacen el requisito de máxima especificidad, con todo, no satisficieran estrictamente el requisito de los elementos de juicio totales en la forma cuantitativa precisa que estamos considerando. Por ejemplo, supongamos que K consista en las premisas de (3o) y el enunciado adicional "Hb"; entonces, aunque intuitivamente este último enunciado sea por completo ajeno a la conclusión "Gb", es concebible que la probabilidad lógica, en el sentido aquí supuesto, de "Gb" relativa a K diferiría de la probabilidad lógica r de "Gb" referida a las premisas de (3o) solamente. O supongamos que K consiste en los enunciados "$p(G, F) = 0{,}9$", "$p(G, H) = 0{,}1$", "$p(G, F \cdot H) = 0{,}85$", "Fb", "Hb"; entonces un razonamiento estadístico con los últimos tres enunciados como premisas y "Gb" como conclusión satisfaría el requisito de especificidad máxima referida a K. Sin embargo, nuevamente la probabilidad lógica de "Gb" referida a K diferiría de la probabilidad lógica, 0,85, de "Gb" referida al conjunto de los tres enunciados-premisas.

Pero en ausencia de una adecuada definición general de probabilidad lógica, parece claro que el razonamiento predictivo que acabamos de analizar sería considerado racionalmente aceptable en la situación cognoscitiva representada por K; la ley estadística que especifica la probabilidad de G en $F \cdot H$ sería considerada como supeditando a sí misma las leyes que especifican la probabilidad de G referida a F y a H, respectivamente. De manera similar, un razonamiento del tipo (3o), cuyas premisas están bien sustentadas y que se ajusta al requisito de la máxima especificidad, seguramente sería considerado como una manera racional de abrigar expectativas concernientes al hecho descripto por la conclusión. Y en general, los razonamientos predictivos de la ciencia que se basan en leyes probabilísticas parecen estar gobernados por el requisito de la máxima especificidad y el de la adecuada confirmación de las premisas. En la medida así indicada, pues, un razonamiento que constituya una explicación estadística aceptable referida a K también constituirá una predicción potencial aceptable referida a K.

Hanson[80] ha expuesto una interesante concepción de la relación entre razonamientos explicativos y razonamientos predictivos en la ciencia que me brinda la ocasión de ampliar la posición general que acabo de esbozar y de enriquecer la argumentación en su favor.

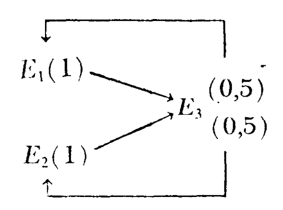
Según Hanson, la idea de que una explicación adecuada también suministra una predicción potencial concuerda con el carácter de las explicaciones y predicciones que permite la mecánica clásica newtoniana, la cual tiene un carácter determinista; pero es totalmente inadecuada con respecto a la teoría cuántica, que es fundamentalmente no determinista. Con más especificidad, Hanson sosiene que las leyes de la teoría cuántica no permiten la *predicción* de ningún fenómeno cuántico individual F, como la emisión de una partícula beta por una sustancia radiactiva, pero que "F puede ser completamente *explicada ex post facto*; podemos comprender cabalmente qué tipo de suceso se produjo en función de las leyes bien estable-

[80] Hanson (1959 y 1963, capítulo 2).

cidas de ... la teoría cuántica ... Estas leyes fijan el *significado* de 'explicar microhechos aislados' ".[81]

Sucede, en verdad, que a causa de su carácter puramente estadístico las leyes de la desintegración radiactiva permiten la predicción de hechos tales como la emisión de partículas beta por átomos en desintegración sólo con probabilidad, y no con precisión deductivo-nomológica, para un hecho individual. Pero por la misma razón, exactamente, esas leyes sólo permiten una explicación probabilística de una emisión particular F, y no una explicación "completa *ex post facto*", como dice Hanson. Pues si la información de que F ha sucedido fuera incluida en el explanans, como la expresión "*ex post facto*" parecería sugerir, la explicación resultante sería poco esclarecedora y circular; y sin duda no es esto lo que Hanson quiere significar. Y si el explanans sólo contiene enunciados acerca de condiciones antecedentes, más las leyes estadísticas de la desintegración radiactiva, entonces sólo puede establecer en el mejor de los casos que F era altamente probable; pero esto sólo brinda una explicación inductivo-estadística que tiene la misma forma lógica que la predicción probabilística —es decir, inductivo-estadística— de F.[82]

En el contexto de este razonamiento Hanson hace otra afirmación, a saber: "Toda predicción, para ser inferencialmente respetable, debe poseer una posdicción correspondiente".[83] Por posdicción Hanson entiende "simplemente la inversión lógica de una predicción": una predicción pasa "de condiciones iniciales a través de condiciones limitantes a un enunciado acerca de algún hecho futuro x", y una posdicción consiste "en inferir de un enunciado acerca de algún suceso actual x, a través de las condiciones iniciales ya *conocidas*".[84] Pero la tesis de Hanson es incorrecta, como lo revela el siguiente contraejemplo. Consideremos un sistema de estados discretos cuyos tres estados posibles E_1, E_2 y E_3 están vinculados por las leyes siguientes: E_3 sigue siempre a E_1 y E_2; E_3 es seguido, con una probabilidad de 0,5 por E_1, y con la misma probabilidad por E_2. El diagrama de transición correspondiente es:



[81] Hanson (1959, pág. 354); bastardillas del autor; lo mismo en Hanson (1963, pág. 29).

[82] Comentarios hechos en un espíritu semejante se hallarán en Hanson (1963); véase también la respuesta crítica en Feyerabend (1964).

[83] Hanson (1963, pág. 193; véase también la pág. 40). Hanson agrega: "Esto es parte de la tesis de Hempel, y es correcto, necesariamente" (*ibíd.*). En realidad, he argüido *en contra* de esa tesis, que es exacta en lo que se refiere a las predicciones basadas en teorías deterministas, pero no de un modo general. Véase Hempel (1962, págs. 114-115).

[84] Hanson (1963, pág. 193); bastardillas del autor.

Entonces, la información de que en el intervalo de tiempo t_5 el sistema está en E_2 permite la predicción deductivo-nomológica y, por ende, "inferencialmente respetable" sin duda, de que durante t_6 el sistema estará en E_3; pero no es posible ninguna posdicción correspondiente que pase de esta última información a la primera.[85]

Para concluir consideraré un argumento expuesto por Rescher sobre la relación entre explicación y predicción. La manera más simple de formular el quid de este argumento es haciendo particular referencia al estudio de Rescher sobre sistemas de estados discretos, al que ya nos referimos en la sección 3.4. Según la definición de Rescher, un razonamiento que explique el estado de tal sistema en el intervalo t puede referirse, en el explanans, a los estados que presenta el sistema en otros momentos, que pueden ser anteriores o posteriores a t; mientras que un razonamiento que prediga el estado del sistema en t sólo debe referirse a estados anteriores. Como consecuencia de estas estipulaciones definicionales, "se sigue que cuando se da ... una predicción, ésta es a fortiori una explicación", pero no a la inversa. "Pues nuestras condiciones definitorias de la predicción ..., en efecto, agregan a las condiciones de la explicación ... ciertas restricciones de carácter temporal".[86]

En pro de la imposición de este requisito adicional a la predicción Rescher arguye del siguiente modo: supongamos que las premisas de un razonamiento propuesto que prediga el estado del sistema en t incluyan un enunciado que especifique el estado de este sistema en un tiempo posterior t_1; entonces, puesto que el razonamiento es predictivo, t es posterior "al presente", t_N, y luego lo mismo sucede con t_1. Ahora bien, existen dos posibilidades. O bien (I) la premisa concerniente a t_1 puede inferirse, por medio de leyes, a partir de estados pasados del sistema, en cuyo caso el razonamiento predictivo dado puede ser reemplazado, evidentemente, por otro que infiera el estado del sistema en t, mediante leyes, exclusivamente a partir de estados pasados, de modo que satisfaga el requisito restrictivo; o bien (II) la premisa explicativa acerca de t_1 no puede inferirse a partir de enunciados acerca de estados pasados: en este caso, "en realidad no tenemos en absoluto una predicción propiamente dicha, pues estamos basando nuestro razonamiento 'predictivo' en una premisa que no es posible justificar por la información *disponible*".[87]

Pero, como indica la referencia a elementos de juicio justificativos, esta consideración no alcanza a la tesis de que un razonamiento explicativo es también potencialmente un razonamiento predictivo, es decir, que se lo podría haber usado para deducir una sentencia predictiva concerniente al estado del sistema en t si se hubieran formulado y utilizado como premisas los enunciados que forman el explanans antes de t. Sin duda, normalmente sólo buscaríamos una explicación de un estado después de su aparición, es decir, en nuestro caso, después de t;[88] y es verdad, como señala el

[85] Sobre este punto véase también Grünbaum (1963, pág. 76). El artículo de Grünbaum ofrece un detallado examen de la identidad estructural de la explicación y la predicción, y examina varias objeciones a esta idea.

[86] Rescher (1963, pág. 329).

[87] Rescher (1963, pág. 333); bastardillas del autor.

[88] En verdad, por un razonamiento semejante, Rescher parece obligado a decir

argumento, que entonces podemos estar en condiciones de sustentar la premisa crítica mediante elementos de juicio de los que no se disponía antes de *t*. Pero el sustento empírico de las premisas no tiene vinculación alguna con las relaciones estructurales entre razonamientos explicativos y razonamientos predictivos; ni creo que las consideraciones basadas en él suministren buenas razones para imponer una condición formal restrictiva a las inferencias predictivas.

Debe recordarse también que, como observamos en la sección 2.4, hasta en los casos más perfectos de predicción científica normalmente se utilizan algunos enunciados acerca del futuro que no se infieren mediante leyes de datos acerca del pasado. Así, la predicción de las posiciones de los planetas en un momento dado sobre la base de los datos requeridos concernientes a sus ubicaciones y cantidades de movimiento un mes antes, exige una suposición relativa a las condiciones limitantes durante el lapso intermedio, habitualmente para excluir la posibilidad de que se hayan producido interferencias externas en el sistema. Y aunque esto no se infiere mediante leyes de otros sucesos particulares, los razonamientos que presuponen esas condiciones limitantes no son considerados por ello como si brindaran alguna predicción propiamente dicha.

Por último, podemos observar con Scheffler que a veces podemos hablar, de manera razonable, de explicar un hecho futuro y que en algunos casos puede considerarse que el mismo razonamiento predice cierto hecho y, al mismo tiempo, lo explica; por ejemplo, cuando se contesta a la pregunta "¿por qué saldrá el sol mañana?", ofreciendo alguna información astronómica adecuada.[89] Por esta razón, también, no parece conveniente imponer requisitos formales diferentes a los razonamientos explicativos y a los predictivos.

3.6. *La ausencia de conjunción de la explicación inductivo-estadística*

La explicación inductivo-estadística difiere de la deductiva en otro aspecto importante. Cuando un explanans dado explica deductivamente cada uno de varios explananda, entonces también explica deductivamente su conjunción; pero esto no es válido en general para la explicación I-E, porque un explanans que confiere una alta probabilidad a cada uno de varios explananda puede conferir una probabilidad muy baja a su conjunción. En este sentido, pues, *la explicación I-E, en contraste con la explicación deductiva, es no-conjuntiva.*

Consideremos, por ejemplo, el experimento de azar *A* consistente en arrojar una moneda equilibrada diez veces. Cada ejecución de este experimento dará como resultado una de las $2^{10} = 1024$ sucesiones posibles diferentes de 10 resultados particulares, cada uno de los cuales es cara o ceca.

que el razonamiento considerado en nuestro ejemplo, una de cuyas premisas se refiere a t_1, tampoco es una explicación si se lo presenta antes de t_1 (pero después de t), pues se basa entonces en una premisa que no se halla justificada por los elementos de juicio disponibles.

[89] Scheffler (1957, pág. 300).

Sean O_1, O_2, ..., O_{1024} los diferentes tipos posibles de resultados caracterizados de ese modo. Entonces, según la hipótesis estadística común —llamémosla E— para este tipo de experimentos, la probabilidad de obtener cara arrojando la moneda es de 1/2, y los resultados de diferentes lanzamientos son estadísticamente independientes unos de otros. Por ende, se sigue deductivamente de ello que la probabilidad estadística de obtener el resultado O_k en la ejecución de es $p(O_k, A) = 1/1024$, y la probabilidad de obtener un resultado distinto de O_k es $p(\overline{O}_k, A) = 1 - 1/1024 = 1023/1024$ para cualquiera de los diferentes resultados posibles O_k.

Supongamos ahora que un caso particular, a, de A ha dado el resultado O_{500} : $O_{500}(a)$. También es posible describir este resultado diciendo que a no dio ninguno de los otros resultados posibles:

$$\overline{O}_1(a) \cdot \overline{O}_2(a) \ldots \overline{O}_{499}(a) \cdot \overline{O}_{501}(a) \cdot \ldots \cdot \overline{O}_{1024}(a)$$

Ahora bien, nuestra hipótesis estadística E junto con la información de que a fue un caso particular de A, es decir que $A(a)$, suministra una explicación I-E con una elevada probabilidad asociada a ella de (los hechos descriptos por) cada una de las 1023 oraciones aquí unidas del siguiente modo:

$$\begin{array}{ll} p(\overline{O}_k, A) = 1023/1024 & \\ A(a) & \\ \hline\hline & [1023/1024] \\ \overline{O}_k(a) & \end{array}$$

Estas explicaciones satisfacen el requisito de la máxima especificidad, ya que para tal información adicional acerca del experimento particular a de la que podemos disponer en estas circunstancias, puede considerarse que E implica que dicha información no influye sobre la probabilidad de $\overline{O}_k$. Pero si bien E en combinación con la información de que $A(a)$ confiere una alta probabilidad a cada uno de los 1023 enunciados unidos por la conjunción, en cambio asigna la bajísima probabilidad de 1/1024 a su conjunción, que es equivalente al enunciado "$O_{500}(a)$".

Pues, tenemos

$$\begin{array}{ll} p(O_{500}, A) = 1/1024 & \\ A(a) & \\ \hline\hline & [1/1024] \\ O_{500}(a) & \end{array}$$

Así, aunque E junto con "$A(a)$" suministra una explicación I-E con una alta probabilidad asociada de (los hechos descriptos por) cualquiera de las 1023 oraciones citadas, en cambio no lo hace para ((los hechos descriptos por) su conjunción.[90]

[90] Se hallará otro ejemplo en Hempel (1962, pág. 165).

Esta no-conjuntividad de la explicación I-E surge, pues, del hecho de que uno y el mismo conjunto de oraciones puede confirmar en un grado muy alto cada uno de *n* enunciados alternativos, pero confirma con fuerza similar la negación de su conjunción. Este hecho, a su vez, se basa en el teorema general de las probabilidades compuestas, el cual implica que la probabilidad de la conjunción de dos elementos (esto es, características u oraciones, según que se trate de probabilidades estadísticas o probabilidades lógicas) es, en general, menor que la probabilidad de uno cualquiera de los elementos tomado aisladamente. Por consiguiente, una vez que se considera como inductiva la conexión entre el explanans y el explanandum de la explicación estadística de fenómenos particulares, la no-conjuntividad se presenta como un aspecto inevitable de ella y, por ende, como una de las características fundamentales que distinguen la explicación I-E de las explicaciones deductivas.

4. *Los conceptos de la explicación por leyes inclusivas como modelos explicativos*

4.1. *Carácter general y propósito de los modelos*

Hemos distinguido tres tipos básicos de explicación científica: la deductivo-nomológica, la inductivo-estadística y la deductivo-estadística. A la primera se la llama a menudo el modelo de ley inclusiva o modelo deductivo de explicación, pero puesto que también los otros dos tipos incluyen una referencia a leyes inclusivas y puesto que uno de ellos es también deductivo, llamaremos al primero más específicamente el modelo *deductivo-nomológico*; de manera análoga, llamaremos a los otros el *modelo de explicación inductivo-estadístico* y el *deductivo-estadístico.*

Como lo aclaran nuestros exámenes anteriores, estos modelos no pretenden describir la manera como los científicos formulan realmente sus descripciones explicativas. Su propósito es más bien indicar en términos razonablemente precisos la estructura lógica y la justificación de diversos modos en que la ciencia empírica responde a las preguntas que piden una explicación. La construcción de nuestros modelos, pues, supone cierto grado de abstracción y de esquematización lógica.

A estos respectos, nuestros conceptos de explicación se asemejan al concepto, o a los conceptos, de prueba matemática (dentro de una teoría matemática dada) tal como se la concibe en la matemática. Observemos los puntos principales de semejanza.

En ambos casos, los modelos tratan de explicar el uso y la función de ciertos términos del "explicandum": "prueba" y sus afines en un caso, "explicación" y sus afines en el otro. Pero los modelos son selectivos; no pretenden aclarar todos los diferentes sentidos habituales de los términos en cuestión, sino solamente algunos sentidos especiales. Así, la teoría metamatemática de la prueba sólo se ocupa de la noción de prueba en la matemática. La exposición de la teoría no equivale a negar que haya otros contextos en los que hablamos de pruebas y de probar, ni a afirmar que los conceptos metamatemáticos sean aplicables a esos contextos.

Análogamente, exponer los modelos por leyes inclusivas de la explicación científica no equivale a negar que haya otros contextos en los cuales hablamos de explicación, ni a afirmar que los sentidos correspondientes de la palabra "explicar" se ajusten a uno u otro de nuestros modelos. Obviamente, tales modelos no pretenden reflejar los diversos sentidos de "explicar" que intervienen cuando hablamos de explicar las reglas de un torneo, explicar el significado de una inscripción cuneiforme, de una cláusula legal compleja o de un pasaje de un poema simbolista, de explicar cómo cocinar un pastel o cómo reparar una radio. Elucidar el concepto de explicación científica no es lo mismo que escribir una nota sobre la palabra "explicar" destina al *Oxford English Dictionary*. Por consiguiente, lamentar como lo hace un crítico la "inutilidad" del modelo deductivo-nomológico sobre la base de que no se ajusta al caso de explicar o comprender las reglas de la sucesión hanoveriana[91] supone simplemente equivocarse con respecto al propósito del modelo. Y es el colmo de la equivocación señalar que el modelo deductivo-nomológico presupone que las explicaciones están formuladas en un "lenguaje descriptivo", cuando "hay casos, obviamente, en los que podemos explicar sin lenguaje, por ejemplo, cuando explicamos al mecánico de un garaje yugoslavo qué es lo que anda mal en nuestro auto".[92] Esto es lo mismo que objetar una definición metamatemática de prueba sobre la base de que no se ajusta al sentido de la palabra "prueba" en "la prueba del budín se hace comiéndolo" o en "whisky de prueba 86". La gesticulación sin palabras tendiente a indicar a un mecánico yugoslavo qué es lo que anda mal en nuestro auto, en efecto, no queda calificada como explicación científica según ninguno de nuestros modelos; pero esto debe ser así, pues una concepción de la explicación científica que admitiera este caso sería seriamente inadecuada.

En apoyo de la idea de que un análisis adecuado de la explicación debe abarcar todos estos sentidos diferentes de la palabra "explicar", Scriven ha sostenido que todos ellos tienen la misma "función lógica", acerca de la cual observa: "La solicitud de una explicación presupone que se comprende *algo,* y una respuesta completa es aquella que relaciona el objeto de la investigación con el ámbito de comprensión de una manera clara y adecuada. Esta manera varía de un tema a otro ...; pero la *función lógica* de la explicación ... es la misma en cada campo".[93] Pero si bien la observación inicial de este pasaje bien puede aplicarse a muy diferentes tipos de explicación, ni ella ni las restantes observaciones de Scriven sobre la cuestión incumben a lo que podría llamarse con propiedad un aspecto *lógico* de la explicación. En verdad, expresiones tales como "ámbito de comprensión" y "clara" no pertenecen al vocabulario de la lógica, pues se refieren a los aspectos psicológicos o pragmáticos de la explicación. Consideraremos estos aspectos en la sección siguiente y veremos que

91 Scriven (1959, pág. 452).
92 Scriven (1962, pág. 192). También Brodbeck (1962, pág. 240) ha señalado que tales objeciones son ajenas a la cuestión. Algunos comentarios agudos y estimulantes sobre este problema o sobre otros aspectos de "la querella acerca de la explicación histórica" se encontrarán en el artículo de Weingartner (1961).
93 Scriven (1962, pág. 202); bastardillas del autor.

cuando se los concibe como observaciones acerca de la pragmática de la explicación más que acerca de su lógica, las caracterizaciones como las de Scriven son totalmente ajenas a la cuestión.

Pero no puede decirse que las diferentes maneras de explicar consideradas por Scriven tienen la misma función lógica. Pues, en primer término, aun los medios lingüísticos que sirven para indicar el tema de los diferentes tipos de explicación tienen un carácter lógico diferente. Por ejemplo, cuando una explicación debe indicar el "significado" de un pasaje literario, un símbolo, una obra de arte, etc., el explanandum estará especificado por medio de una *expresión sustantiva* ("el signo de interrogación", "la primera frase del Génesis", "la cruz gamada"); mientras que las explicaciones que hemos estado considerando se refieren a hechos, fenómenos, sucesos, uniformidades, etc., cada uno de los cuales queda caracterizado adecuadamente por medio de una *oración* (que aparece como la oración-explanandum en nuestros esquemas). En segundo término, el problema de especificar significados y el de enunciar las "causas" de un fenómeno o las razones por las cuales se emprendió una acción tienen, sin duda, un carácter lógico diferente; y la adecuación de las soluciones propuestas en cada caso debe juzgarse con criterios muy diferentes. Las diferencias entre las tareas que deben cumplir estos y otros tipos de explicación son, de hecho, precisamente, diferencias entre la estructura lógica de los tipos correspondientes de explicación.

Pasemos ahora de la selectividad de los modelos elucidadores de las pruebas y de la explicación a otro rasgo común. La teoría metamatemática de la prueba no pretende dar una explicación descriptiva de la manera como los matemáticos formulan sus pruebas. En realidad, las formulaciones que los matemáticos ofrecen habitualmente se apartan en cierta medida de los patrones metamatemáticos rigurosos y, por así decir, "ideales". Sin embargo, puede decirse que esos patrones exponen la estructura lógica y la justificación de la demostración matemática, y brindan criterios para la estimación crítica de las pruebas particulares que puedan proponerse.

Puede hallarse, pues, que una prueba propuesta se aparta de una pauta teórica determinada sólo en aspectos poco importantes; por ejemplo, omitiendo ciertos pasos intermedios del razonamiento por considerarlos obvios, o no mencionando ciertas premisas, que se dan por supuestas pero que pueden ser especificadas de manera explícita si ello es necesario. En tales casos, podemos decir que la prueba está *formulada elípticamente.* Por otra parte, los defectos pueden ser fundamentales, como en las diversas pruebas del postulado de las paralelas sobre la base de los otros postulados de la geometría euclidiana.

Además de brindar normas para la estimación crítica, la construcción de conceptos rigurosos de prueba matemática ha permitido la elaboración de una teoría poderosa que ha dado resultados de largo alcance y a menudo totalmente inesperados concernientes a la demostrabilidad, la decidibilidad y la definibilidad en sistemas matemáticos de tipos específicos.

Los modelos analíticos de explicación científica, creo yo, pueden servir a propósitos similares, aunque en una escala mucho más modesta. En cuanto a la posibilidad de desarrollos sistemáticos generales podemos men-

cionar, por ejemplo, los resultados establecidos por Ramsey y Craig[94] acerca del papel y la posible prescindibilidad, en el contexto de la explicación científica, de principios que se refieren ostensiblemente a entidades "teóricas" inobservables. Sólo es posible lograr estos resultados, y la comprensión que brindan de la lógica del procedimiento científico, con referencia a una concepción precisamente formulada, y en cierta medida esquemática, de la explicación científica.

4.2. *Variedades de incompletividad explicativa*

4.2.1. *Formulación elíptica.* Al igual que una prueba matemática propuesta, una explicación propuesta puede estar *formulada elípticamente.* Por ejemplo, cuando explicamos que un trozo de manteca se derritió porque fue colocado en una sartén caliente, o que apareció un pequeño arco iris en el rocío de una regadera porque la luz del sol se reflejó y refractó en las gotitas de agua, puede decirse que ofrecemos versiones elípticas de explicaciones D-N. Las explicaciones de este tipo no mencionan ciertas leyes o hechos particulares que se dan tácitamente por supuestos y cuya inclusión explícita en el explanans brindaría un razonamiento D-N completo. Puede decirse que una explicación formulada elípticamente es *incompleta,* pero en un sentido más bien inocuo.

4.2.2. *Explicación parcial.* Pero a menudo las exposiciones explicativas manifiestan un tipo más serio de incompletividad. En este caso, los enunciados realmente incluidos en el explanans, aun completados con aquellos respecto de los cuales puede suponerse razonablemente que se los da por supuesto de manera tácita en el contexto dado, sólo explican parcialmente el explanandum especificado, en un sentido que trataré de aclarar mediante un ejemplo.

En su obra *Psicopatología de la vida cotidiana,* Freud ofrece esta descripción y explicación de un lapsus al escribir:

> En una hoja de papel que contenía principalmente breves notas cotidianas, hallé con sorpresa la fecha incorrecta "jueves, octubre 20" colocada entre corchetes debajo de la fecha correcta del mes de septiembre. No fue difícil explicar esta anticipación como la expresión de un deseo. Pocos días antes había vuelto de mis vacaciones y sentía grandes deseos de realizar una intensa labor profesional, pero todavía había pocos pacientes. A mi llegada, encontré una carta de un paciente que me anunciaba su llegada para el 20 de octubre. Como escribí la misma fecha en septiembre, ciertamente debo de haber pensado: "X ya debería estar aquí; ¡qué lástima todo un mes perdido!" Y con este pensamiento, corrí la fecha un mes después.[95]

Evidentemente, esta formulación de la explicación es elíptica en el sentido considerado antes; pues no menciona leyes o principios teóricos en

[94] Véase Ramsey (1931, págs. 212-215, 231); y Craig (1956). Véase también el examen de estos resultados en Hempel (1958, sección 9).
[95] Freud (1951, pág. 64).

virtud de los cuales podría sostenerse que el deseo subconsciente y las otras circunstancias particulares aludidas explican el desliz en cuestión. Pero las ideas teóricas que propone Freud para la interpretación de tales deslices sugieren claramente que su explicación está regida por una hipótesis general según la cual cuando una persona tiene un fuerte deseo, aunque quizá subconsciente, si incurre en un desliz al escribir, hablar o recordar, éste adoptará una forma que exprese —y quizá satisfaga— simbólicamente ese deseo.

Aun esta vaga desclaración es, sin duda, más definida que lo que Freud habría estado dispuesto a afirmar; y quizás, a pesar de las propensiones deterministas de Freud, sería más adecuado concebir la hipótesis fundamental como de forma estadística, y considerar como probabilística la explicación propuesta. Pero para servir a los fines de nuestra argumentación, aceptemos la hipótesis tal como la hemos formulado e incorporémosla al explanans, junto con enunciados particulares que afirmen que Freud tenía el deseo subconsciente que menciona, y que de hecho iba a cometer un desliz de la pluma. Aun en tal caso, el explanans resultante sólo nos permite inferir que el desliz adoptará *una u otra forma* que expresará, y quizá satisfará simbólicamente, el deseo subconsciente de Freud; pero el explanans no implica que el desliz adoptará la forma específica de escribir "jueves, octubre 20" en el calenario, junto a la fecha correspondiente de septiembre.

Pero en la medida en que la clase *F* de desliz que adopta esta forma sea una subclase propia de la clase *W* de los deslices al escribir que expresen de alguna manera y quizá satisfagan simbólicamente el deseo especificado, podemos decir que el explanandum, tal como lo describe Freud —es decir, que cometió un desliz que pertenece a la clase *F*—, queda explicado al menos en parte por su explicación, que coloca el desliz en la clase más amplia *W*. Los razonamientos de este tipo pueden llamarse *explicaciones parciales*. Muchas de las explicaciones que se ofrecen en el psicoanálisis[96] y en la historiografía son, a lo sumo, explicaciones parciales en este sentido: el explanans no explica el fenómeno expresado en el explanandum con la especificidad que caracteriza a la oración explanandum, por lo cual la fuerza explicativa del razonamiento es menor de lo que pretende o parece ser.

Creo que es importante y esclarecedor distinguir tales explicaciones parciales, por grande que sea la frecuencia con la que se brindan y se aceptan y por fecundas y sugestivas que sean, de lo que podría llamarse *explicaciones deductivamente completas*, es decir, aquellas en las cuales el explanandum, tal como está formulado, es implicado lógicamente por el explanans; pues estas últimas explican el fenómeno aludido en el explanandum con la especificidad con la que la oración del explanandum lo describe.[97]

[96] Esto es válido, creo, para los muchos, y a menudo sumamente sugerentes, análisis que se encuentran en la *Psicopatología de la vida cotidiana* de Freud.

[97] Una explicación parcial puede ser, evidentemente, más o menos débil, según la medida en que la clase dentro de la cual ubica el explanans el caso dado (*W* en nuestro ejemplo) sea más extensa que la clase a la cual lo asigna la oración-explanandum (*F* en nuestro caso). Además, mientras que algunas explicaciones par-

Una explicación que se ajusta al modelo D-N, pues, es automáticamente completa en este sentido; y una explicación parcial como la que hemos caracterizado nunca es una explicación D-N.

En la explicación estadística, el explanans no implica lógicamente el explanandum. ¿Debemos calificar entonces a todas esas explicaciones como incompletas? Dray plantea esta cuestión cuando pregunta si "un hecho puede explicarse *completamente* (aunque quizás en un sentido diferente) sin subsumirlo bajo una ley universal que autorice su deducción, y por consiguiente sin mostrar que debía suceder".[98] La respuesta de que las explicaciones estadísticas son deductivamente incompletas sería una trivialidad carente de interés. Como lo sugiere la aclaración de Dray "aunque quizás en un sentido diferente", estamos más bien ante la cuestión de si la noción de completividad explicativa, definida hasta ahora solamente con referencias a las explicaciones D-N propuestas, podría ampliarse razonablemente hasta ser aplicable asimismo dentro del dominio de la explicación probabilística. No parece conveniente construir un concepto ampliado de completividad explicativa para calificar de incompletas a todas las explicaciones estadísticas. Pues esta calificación tiene una connotación de deficiencia y, sin duda, no podemos considerar a las explicaciones estadísticas simplemente como explicaciones D-N fracasadas: constituyen por sí mismas un tipo importante de explicación. Indudablemente, los primeros usos explicativos de leyes y teorías estadísticas por ejemplo, en las física del siglo XIX, a menudo se inspiraban en la creencia de que los microfenómenos implicados en los procesos físicos en estudio estaban todos sujetos a leyes estrictamente universales, y que el uso de hipótesis y teorías estadísticas era necesario solamente por las limitaciones de nuestra capacidad individual para medir todos esos microfenómenos, y luego realizar los grandes y complejos cálculos que se requerían para explicar un fenómeno físico con todo detalle microscópico. Pero esta idea ha sido abandonada gradualmente: en ciertas ramas de la física, como la teoría cuántica, se han aceptado las leyes de forma estadística como leyes básicas de la naturaleza. Y cualquiera que sea el futuro de la teorizacón científica, este desarrollo refleja claramente la comprensión de que la explicación estadística es por lógica independiente de la suposición de leyes universales estrictas y constituye un modo de explicación sui géneris. Todo esto sugiere claramente

---

ciales son, sin duda, esclarecedoras y sugieren nuevas investigaciones que pueden conducir a una explicación más completa, hay otros razonamientos que carecen totalmente de tal mérito, aunque presenten una semejanza formal con nuestro ejemplo y puedan, por esta razón, ser calificados como explicaciones parciales. Supongamos, por ejemplo, que *b* es *F* y también *G*, y que tenemos una explicación D-N de que *b* sea *F*. Entonces (salvo para ciertas excepciones triviales), el explanans de la última suministrará automáticamente una base para una explicación parcial de que *b* sea *G*, pues implica que *b* es *F* y, por tanto, que *b* es *F* o *G*, y la clase caracterizada por "*F* o *G*" contiene a *G* como subclase propia. Pero no es mi propósito aquí explorar las condiciones en las cuales las explicaciones parciales pueden resultar útiles; simplemente quiero llamar la atención sobre el hecho de que muchas explicaciones que se ofrecen en la bibliografía de las ciencias empíricas tienen las características formales de las explicaciones parciales y que, en consecuencia, exageran la medida en la cual explican un fenómeno dado.

[98] Dray (1963, pág. 119).

que en una razonable extensión de la idea de completividad explicativa, toda explicación que se conforme a nuestro modelo estadístico debe ser formalmente completa, pues asigna al suceso descripto por el enunciado del explanandum (o, más propiamente, al enunciado mismo del explanandum) la probabilidad lógica que exige la relación lógica entre los enunciados del explanans y del explanandum. A este respecto, tal explicación estadística es análoga a la que se ajusta al modelo D-N, por lo cual afirma correctamente que el explanandum está implicado en el explanans (y, por consiguiente, tiene la probabilidad lógica 1 con respecto a éste). A la luz de esta analogía una explicación estadística propuesta debe calificarse como parcial si el explanans confiere la probabilidad especificada, no a la oración del explanandum realmente enunciada, sino a otra más débil relacionada con ella de la manera que ilustra nuestro ejemplo tomado de Freud. Podemos ilustrar de modo muy esquemático la idea con referencia a este mismo ejemplo. Supongamos que concibiéramos la ley general que formulamos tentativamente como base presuntiva de la explicación de Freud igual que una ley estadística según la cual en presencia de un fuerte deseo, aunque subconsciente, se eleva la probabilidad de que si se comete un desliz de la pluma, adoptará una forma que exprese, y quizá satisfaga, simbólicamente este deseo. Luego la explicación de Freud —concebida ahora como afirmando que la información explicativa aducida confiere una alta probabilidad lógica al enunciado del explanandum— sería un *explicación estadística parcial*; pues el explanans confiere una elevada probabilidad, no al anunciado de que el desliz particular cae dentro de la clase *F* definida antes, sino al enunciado más débil de que el desliz pertenece a la clase *W*.

4.2.3. *Incompletividad explicativa y sobredeterminación.* Las consideraciones que acabamos de hacer son también atinentes al problema ilustrado por el siguiente ejemplo:[99] Supongamos que la barra *b*, hecha de cobre (*Cb*), es sometida simultáneamente al calor (*Hb*) y a tensión longitudinal (*Tb*), y que en el proceso la barra se alarga (*Ab*). Entonces es posible formular dos razonamientos diferentes, cada uno de los cuales, según las patrones que hemos sugerido, constituye una explicación D-N de por qué la barra se alargó. Uno de ellos se basará en la ley de que las barras de cobre se alargan cuando se calientan; el otro, en la ley de que las barras de cobre se alargan cuando se las estira. Esquemáticamente:

$$\frac{(x)\ [(Cx \cdot Hx) \supset Ax] \qquad Cb \cdot Hb}{Ab} \qquad\qquad \frac{(x)\ [(Cx \cdot Tx) \supset Ax] \qquad Cb \cdot Tb}{Ab}$$

[99] Debo a mi colega en Princeton, el profesor Arthur Mendel, del Departamento de Música, el haberme planteado algunas cuestiones acuciantes que me hicieron tomar conciencia del problema aquí considerado. En su artículo (1962), Mendel toma como punto de partida un problema concreto de la historia de la música y, con referencia a éste, desarrolla algunas aclaradoras ideas concernientes, entre

Podría objetarse que aun admitiendo la verdad de todas las premisas, ambas explicaciones son inaceptables por "incompletas": cada una de ellas ignora uno de los dos factores que contribuyeron al alargamiento. Para estimar la fuerza de esta objeción es, nuevamente, importante comprender con claridad qué es lo que se debe explicar. Si se trata simplemente, como en nuestro ejemplo, de explicar el hecho *Ab*, es decir, que *b* se alargó o que hubo *un* aumento en la longitud de *b*, entonces, creo, cualquiera de los dos razonamientos lo explica de manera concluyente, y la acusación de incompletividad es infundada. Pero si deseamos explicar el hecho de que la longitud de la barra aumentó en tanto y tanto, entonces, evidentemente ninguno de los dos razonamientos lo explica; pues tendríamos que tomar en consideración tanto el aumento de la temperatura como el estiramiento, y necesitaríamos leyes cuantitativas referentes a su efecto conjunto sobre la longitud de una barra de cobre. Es necesario administrar con cuidado las locuciones tales como "explicar el aumento de longitud de una barra de metal": ellas son ambiguas en cuanto to se refiere a las dos tareas, al menos, totalmente diferentes que hemos distinguido.

Adoptando un término que se usa con frecuencia en la teoría psicoanalítica podríamos decir que un hecho está sobredeterminado si se dispone de dos o más explicaciones alternativas con conjuntos-explanans no equivalentes. Así, la aparición de un alargamiento en la barra de cobre *b* constituye un caso de ***sobredeterminación explicativa***, ya que se dispone de las explicaciones alternativas mencionadas antes. En este ejemplo, las explicaciones alternativas invocan diferentes leyes (y por ende, algunos enunciados diferentes relativos a hechos particulares). En otro tipo de situación, quizá menos interesante, que sería considerada igualmente como sobredeterminación explicativa según nuestra definición, las explicaciones alternativas se basan en las mismas leyes, pero aducen circunstancias particulares diferentes.[100] Por ejemplo, puede explicarse el estado de un sistema físico determinista en el tiempo *t*, con ayuda de las leyes atinentes al caso, especificando el estado del sistema en cualquier momento anterior; potencialmente, esto permite un número infinito de explicaciones alternativas tales que dos cualesquiera de ellas no tendrían conjuntos explanans lógicamente equivalentes.

Scriven ha planteado un problema que presenta cierta semejanza con el que acabamos de considerar, y lo ilustra mediante el siguiente ejemplo: para explicar cómo fue destruido en tiempo de guerra cierto puente, "apelamos a la ley 'cuando se lanza una bomba atómica exactamente sobre un puente y explota con toda su fuerza, el puente queda destruido', más las condiciones antecedentes adecuadas". Pero también podría "suceder que, cuando se arrojan 1000 kilogramos de dinamita sobre el tramo principal de dicho puente, se lo destruye, y que el movimiento subterráneo ha dado justamente este tratamiento al puente, cuya destrucción se produce entre el lanzamiento y la llegada de la bomba atómica". Scri-

otras cosas, a la significación de los modelos de leyes inclusivas para los objetivos explicativos del historiador.

[100] Sobre este punto, véanse las observaciones de Braithwaite (1953, pág. 320).

ven sostiene que esto invalida la explicación por la bomba atómica, "que no puede dar cuenta de otros aspectos del suceso, en este caso el momento de la destrucción". Y llega a la conclusión de que para excluir tales explicaciones debemos imponer el requisito de que se reúnan elementos de juicio totales, aun en las explicaciones D-N, en una forma más específica, por la cual "una explicación de un fenómeno sólo será aceptable mientras no haya hechos conocidos acerca de la circunstancia de la producción del fenómeno que dicha explicación no pueda acomodar".[101]

Pero sin duda, la explicación por la bomba del ejemplo de Scriven es inaceptable porque su explanans exige la suposición de que, en el lugar en cuestión, había un puente que podría ser destruido por las ondas de presión de la bomba, suposición que es falsa puesto que en ese momento el puente ya había sido destruido por la dinamita. Luego, la explicación por la bomba atómica es falsa en el sentido especificado en la sección II, y no es necesario ningún requisito adicional para desecharla al igual que otras explicaciones de esta especie.

Además de ser innecesario, el requisito específico sugerido por Scriven para descartar la explicación por la bomba atómica y otras similares es demasiado riguroso, creo, para ser sostenible. Pues ni en la investigación científica ni en nuestros asuntos prácticos exigimos una explicación que acomode todo lo que sabemos —o creemos saber— acerca de los hechos que rodean al fenómeno del explanandum. En el caso del puente, por ejemplo, estos hechos pueden incluir una gran cantidad de información acerca de la forma, el tamaño y la ubicación de los fragmentos después de la destrucción; quizá también, la identidad de los dinamiteros, sus objetivos y muchas otras cosas. Sin duda, no exigimos que todos estos detalles sean explicados por toda explicación aceptable de "cómo el puente llegó a ser destruido".

Finalmente, la condición propuesta por Scriven no tiene nada que ver con el requisito de los elementos de juicio totales; en particular, no es una versión "más específica" de aquél. Y la afirmación de Scriven de que alguna condición semejante debe imponerse aun a las explicaciones de forma deductiva, porque ellas no satisfacen automáticamente el requisito de los elementos de juicio totales,[102] pasa por alto la prueba directa de lo contrario.[103]

4.2.4. *Incompletividad explicativa y "sucesos concretos"*. Como observamos antes, puede considerarse una explicación científica como una respuesta potencial a una pregunta de la forma "¿por qué se da el caso

[101] Scriven (1962, págs. 229-230). Véanse también una breve observación que parece tener la misma intención, en Scriven (1963a, págs. 348-349).

[102] Scriven (1962, pág. 230).

[103] En un razonamiento deductivamente válido, las premisas son un fundamento concluyente para afirmar la conclusión; y toda parte de los elementos de juicio totales que no esté incluida en las premisas es ajena a la conclusión, en el sentido estricto de que si se la agrega a las premisas, el conjunto resultante de oraciones seguiría siendo un fundamento concluyente en favor de la conclusión. En la terminología de la lógica que las premisas de un razonamiento D-N asignan a la conclusión es 1; y sigue siendo 1 si se agrega a las premisas parte de todos los elementos de juicio totales.

de que *p*?", donde el lugar de *p* lo ocupa una oración empírica que detalla los hechos que se quiere explicar. Por consiguiente, tanto el modelo de explicación deductivo-nomológico como el estadístico caracterizan el fenómeno del explanandum por medio de una oración, la oración-explanandum. Tomemos, por ejemplo, la explicación de hechos aislados, tales como el de la longitud de determinada barra de cobre *b* que aumenta durante el intervalo comprendido entre las 9,00 y las 9,01 de la mañana, o la de que una extracción *d* particular de una urna determinada dio un bolilla blanca: aquí los fenómenos del explanandum son descriptos totalmente por las oraciones "la longitud de la barra de cobre *b* aumentó entre las 9,00 y las 9,01 de la mañana" y "la extracción *d* dio una bolilla blanca". Y los hechos o sucesos particulares son reducibles a una explicación científica sólo cuando se los entiende en este sentido, como describibles totalmente por medio de oraciones.

Pero la noción de suceso único o particular a menudo se concibe de una manera muy diferente. En este segundo sentido, se especifica un hecho, no por medio de una oración que lo describe, sino de una frase sustantiva, tal como un nombre de individuo o una descripción definida, por ejemplo, "el primer eclipse solar del siglo XX". "La erupción del Vesubio del año 79", "el asesinato de León Trotsky", "la crisis de la bolsa de 1929". Por falta de una terminología mejor, a los sucesos individuales entendidos de este modo los llamaremos *sucesos concretos*,[104] y a los hechos y sucesos en el primer sentido considerado los llamaremos caracterizables oracionalmente o, más brevemente, *hechos y sucesos oracionales*.

El conocido problema de si es posible dar una explicación completa de sucesos únicos se inspira, sin duda, en la concepción de un suceso único como concreto. Pero, ¿qué podría significar en este caso una explicación completa? Presumiblemente, una explicación que da cuenta de todo aspecto del suceso dado. Si ésta es la idea, entonces ningún suceso concreto puede explicarse completamente, en efecto, pues un suceso concreto tiene infinitos aspectos diferentes y, por ende, ni siquiera se lo puede describir en su totalidad, y mucho menos explicar de igual modo. Por ejemplo, una descripción completa de la erupción del Vesubio del año 79 tendría que especificar el momento exacto en que se produjo; el camino que siguió la corriente de lava, las características físicas y químicas de ésta —incluso las temperaturas, presiones, densidades, etc., en cada punto— y sus cambios en el curso del tiempo; los detalles más minuciosos de la destrucción de Pompeya y Herculano; una información completa acerca de todas las personas y animales afectados por la catástrofe, incluso el hecho de que los restos de tales y cuales víctimas, hallados en tales y cuales lugares, se encuentran expuestos en un museo de Nápoles; y así *ad infinitum*. Debemos mencionar también, ya que también constituye otro aspecto de este suceso concreto, toda la literatura es-

[104] No quiero insinuar que la noción de suceso concreto aquí aludida sea totalmente clara. En particular, no sé cómo formular una condición necesaria y suficiente de identidad para sucesos concretos. Las agudas observaciones de Gibson sobre "Lo explicado" (1960, págs. 188-190) son importantes para los problemas que examinaremos aquí.

crita al respecto. En verdad, no parece haber en modo alguno una manera clara y satisfactoria de excluir una clase de hechos que no constituyan aspectos del suceso concreto al que nos estamos refiriendo. Es obvio, pues, que no tiene objeto pedir una explicación completa de un suceso particular, entendido de tal modo.

En resumen, sólo tiene sentido pedir una explicación concerniente a lo que hemos llamado hechos y sucesos oracionales; sólo con respecto a ellos podemos plantear una pregunta de la forma: "¿por qué se da el caso de que *p*?" En cuanto a los sucesos concretos, observemos que lo que hemos llamado sus aspectos o características es describible mediante oraciones; cada uno de estos aspectos, pues, es un hecho o suceso oracional (por ejemplo, que la erupción del Vesubio del año 79 duró tantas horas, que mató a más de 1000 personas en Pompeya, etc.). Con referencia a tales aspectos particulares de un suceso concreto, pues, tiene sentido pedir una explicación. Y está claro que cuando hablamos de explicar un suceso particular, como la abdicación de Eduardo VIII, normalmente pensamos sólo en ciertos aspectos del suceso como sujetos a examen; cuáles sean los aspectos así enfocados para dar una explicación de ellos dependerá del contexto de la investigación.[105]

Aunque por lo general los problemas que aquí abordamos suelen discutirse con especial referencia a los sucesos históricos en su "individualidad y singularidad", los problemas inherentes a la noción de suceso concreto no se limitan en modo alguno al dominio del historiador. Un suceso tal como el eclipse solar del 20 de julio de 1963 también posee una infinidad de aspectos físicos, químicos, biológicos, sociológicos y otros aun, y no se presta a una descripción completa ni, *a posteriori*, a una explicación completa. Pero ciertos aspectos del eclipse —tales como su duración total y el hecho de que fuera visible en Alaska y luego en Maine— bien pueden ser susceptibles de explicación.

Sería incorrecto, sin embargo, resumir esto diciendo que el objeto de una explicación es siempre un *tipo* de suceso y no un suceso particular. Un tipo de suceso tendría que ser caracterizado por medio de una expresión-predicado, tal como "eclipse solar total" o "erupción volcánica"; y puesto que este tipo de expresión no es una oración, no tiene sentido pedir una explicación de un tipo de sucesos. Lo que en realidad puede explicarse es la *aparición de un caso particular de un tipo determinado de sucesos*, por ejemplo, la producción de un eclipse solar total el 20 de julio de 1963. Y lo que se explica de este modo es definidamente un suceso particular; en verdad, es un suceso único e irrepetible en virtud de la ubicación temporal que se le asigna. Pero es un suceso particular *oracional*, por supuesto: se lo puede describir por medio del enunciado de que el 20

[105] Como observa Max Weber, con especial referencia a la explicación histórica: "Cuando se dice que la historia trata de comprender causalmente la *realidad* de un 'suceso' en su individualidad, lo que obviamente no se entiende por esto... es que deba ... explicar causalmente la *realidad* concreta de un suceso en la totalidad de sus cualidades individuales. Hacer esto último no sólo sería realmente imposible, sino que también constituiría en principio una tarea carente de significado". (Weber, 1949, pág. 169. Bastardillas del autor.)

de julio de 1963 se produjo un eclipse de sol total. Estoy de acuerdo, pues, con el rechazo de Mandelbaum de la tesis de Hayek, según la cual la explicación y la predicción nunca se refieren a su suceso particular sino siempre a fenómenos de cierto tipo: "Pensaríamos que la predicción de un eclipse solar específico, o la explicación de este eclipse, se refiere a un suceso particular, aunque no se refiera a todos los aspectos del suceso, como la temperatura del sol o el efecto del eclipse sobre la temperatura de la tierra, etcétera." [106]

Sin embargo, a partir de esta noción de explicación de un particular eclipse solar o de un arco iris determinado, etc., se puede hablar, *como derivación,* de una explicación teórica de los eclipses solares o de los arco-iris en general: tal explicación es, pues, la que da cuenta de cualquier caso de eclipse o arco-iris. Luego, la noción de explicación de casos particulares de un tipo dado es la noción primaria.

4.2.5. *Cierre explicativo: esbozo de explicación.* Quizás puede considerarse atinente a la idea de explicación otro concepto de completividad; lo llamaremos el concepto de cierre explicativo. Una explicación sería completa, en este sentido, si para todo hecho o ley aducidos, contiene a su vez una explicación. En una explicación con cierre explicativo nada quedaría inexplicado. Pero, obviamente, la completividad en este sentido exige un regreso infinito en la explicación, que es entonces interminable. Buscar tal completividad, pues, es equivocarse sobre la naturaleza de la explicación.

En toda etapa del desarrollo de la ciencia empírica, ciertos (presuntos) hechos serán inexplicables; en particular, los expresados por las leyes o principios teóricos fundamentales aceptados en ese momento, aquellos para los que no se dispone de ninguna explicación basada en una teoría "más profunda". Pero aunque inexplicados, estos principios supremos no tienen por qué carecer necesariamente de fundamento, pues, al igual que las hipótesis en las ciencias empíricas, deberán ser susceptibles de someterse a prueba, y bien puede suceder que, de hecho, tests adecuados hayan suministrado fuertes elementos de juicio a su favor.

Hemos considerado ya las diversas formas en las que una explicación propuesta puede apartarse de las normas incorporadas a nuestros modelos analíticos. En algunos casos, lo que se presente como explicación divergirá aun más de esas normas. Por ejemplo, una explicación propuesta que no sea suficientemente explícita y específica como para que se la pueda considerar razonablemente como una explicación formulada de manera elíptica o una explicación parcial, a menudo se la podrá entender como un *esbozo de explicación,* esto es, como el esquema general de algo que, por gradual elaboración y complementación, podría convertirse en un razonamiento explicativo desarrollado con más detalle, basado en hipótesis enunciadas de manera más completa y que permitan realizar una estimación crítica con referencia a elementos de juicio empíricos.

La decisión acerca de si una explicación debe considerarse como de-

106 Mandelbaum (1961, pág. 233).

ductivo-nomológica o estadística formulada elípticamente, como una explicación parcial, como un esbozo de explicación o, quizá, de ninguna de estas maneras, es una cuestión de interpretación juiciosa. Exige una apreciación de la intención de la explicación dada y de las suposiciones básicas que puedan haber quedado sin formular porque se supone que se las comprende en el contexto dado. No es posible formular inequívocos criterios de atribución con este fin, como no es posible decidir si un razonamiento enunciado de modo informal y que no satisface las normas razonablemente estrictas de validez deductiva debe considerarse válido —a pesar de todo— pero formulado de modo entimemático, como falaz, como un correcto razonamiento inductivo o, quizá por falta de claridad, como ninguna de estas cosas.

Entre los diversos aspectos considerados en los que una explicación o una demostración propuesta puede apartarse de las normas lógicas incorporadas a un modelo no pragmático de explicación o prueba, hay varios que sólo pueden caracterizarse con referencia al conocimiento, los intereses, las intenciones, etc., de las personas que proponen los razonamientos en cuestión o de aquellos a quienes se dirigen; por consiguiente, los conceptos correspondientes son esencialmente pragmáticos. Esto es verdad, por ejemplo, respecto de las nociones de entimema, de explicación formulada elípticamente y de esbozo de explicación.

4.3 *Observaciones finales sobre los modelos de leyes inclusivas*

Hemos visto, pues, que las explicaciones formuladas realmente en la ciencia y en contextos cotidianos varían mucho en cuanto a la manera explícita, completa y precisa de especificar el explanans y el explanandum; por lo tanto, divergen más o menos marcadamente de los modelos ideales y esquemáticos de leyes inclusivas. Pero, admitido esto, pienso que todas las explicaciones científicas adecuadas y sus análogas cotidianas afirman o presuponen, al menos implícitamente, el carácter subsumible, inductivo o deductivo de todo lo que debe explicarse bajo leyes o principios teóricos generales.[107] En la explicación de un suceso particular, se necesitan esos principios nómicos generales para vincular el suceso del explanandum con otros sucesos particulares, y es mediante tal vinculación nómica como estos últimos adquieren el status de hechos explicativos. En la explicación de regularidades empíricas más generales, los principios nómi-

[107] Es menester distinguir claramente esta idea de otra que yo no propongo, a saber, la de que todo fenómeno empírico puede explicarse por subsunción deductiva o inductiva en leyes inclusivas. La idea sugerida aquí es que la lógica de toda explicación científica es básicamente de la variedad de ley inclusiva, pero no que todos los fenómenos empíricos sean científicamente explicables y, menos aún, por supuesto, que todos ellos estén gobernados por un sistema de leyes deterministas. La cuestión de si todos los fenómenos empíricos pueden explicarse científicamente no es tan inteligible como podría parecer a primera vista, y exige mucha clarificación analítica. Me inclino a pensar que no es posible darle ningún significado claro. Sea como fuere, y hablando en términos muy amplios, no es posible formarse una opinión acerca de cuáles leyes son válidas en la naturaleza y cuáles fenómenos pueden ser explicados sobre bases analíticas solamente, sino que debe basarse en los resultados de la investigación empírica.

cos aducidos expresan uniformidades más amplias, de las que son especializaciones estrictas o aproximadas las uniformidades que se quiere explicar. Y los modelos de leyes inclusivas representan, hasta donde alcanzo a ver, la estructura lógica básica de los modos principales de tal subsunción explicativa.

La concepción que aquí hemos resumido en líneas generales no puede recibir, por supuesto, una "prueba" estricta; debe juzgarse su valor por la luz que puede arrojar sobre la justificación y la fuerza de las explicaciones ofrecidas en diversas ramas de las ciencias empíricas. Al desarrollar los modelos de leyes inclusivas y caracterizar la función que pretendemos darles ya hemos sugerido algunos de los aspectos en que esta concepción de la explicación puede resultar esclarecedora; otros aspectos similares aparecerán más adelante, en particular cuando dirijamos nuestra atención, en secciones posteriores, al análisis de ciertos procedimientos explicativos peculiares que parecen entrar en conflicto con la concepción de la explicación que se basa en leyes inclusivas.

5. *Aspectos pragmáticos de la explicación*

5.1 *Observaciones introductorias*

Hablando en términos muy generales, explicar algo a una persona es hacérselo claro e inteligible, hacérselo comprender. Así concebida, la palabra "explicación" y las palabras afines a ella son términos *pragmáticos*: su uso exige una referencia a las personas que intervienen en el proceso de explicación. En un contexto pragmático podríamos decir, por ejemplo, que una explicación A explica el hecho X a la persona P_1. Debemos tener presente, pues, que la misma explicación puede no constituir una explicación de X para otra persona P_2, quien podría no considerar siquiera que X requiriera una explicación o podría hallar la explicación A ininteligible o poco esclarecedora, o ajena a aquello que lo desconcierta en X. En este sentido pragmático, pues, la explicación es una noción relativa: puede decirse significativamente que algo constituye una explicación, en este sentido, sólo para este o aquel individuo.

De modo muy similar, la palabra "prueba" y las palabras afines a ella pueden usarse en un sentido pragmático que haga referencia a quienes producen y a quienes reciben los razonamientos. Por ejemplo, un razonamiento Y que demuestre un teorema geométrico simple a la total satisfacción de un principiante puede ser por completo inaceptable y, por ende, no construir prueba alguna para un matemático, y recíprocamente, lo que para el matemático es una prueba correcta y esclarecedora puede ser ininteligible o ajena a la cuestión para el principiante. En general, que un razonamiento dado, Y, pruebe (o explique) cierta cuestión X a una persona P depende, no sólo de X e Y, sino también, en importante medida, de las creencias de P en ese momento, de su inteligencia, sus normas críticas, su idiosincrasia personal, etcétera.

Los aspectos pragmáticos de la prueba constituyen un tema intere-

sante e importante de investigación empírica. Piaget, por ejemplo, ha dedicado un gran esfuerzo al estudio psicológico de los estándares de prueba en niños de diferentes edades. Pero, para los fines de la matemática y la lógica como disciplinas objetivas, necesitamos evidentemente un concepto de prueba que no sea subjetivo, en el sentido de que se refiera a los individuos y varíe según éstos; un concepto en función del cual tenga sentido decir que un razonamiento dado, *Y*, es una prueba de una oración dada, *X* (de una teoría), sin hacer mención alguna de las personas que puedan tomar conocimiento de *Y*. Los conceptos de prueba que tengan este carácter pueden definirse una vez que se ha formalizado adecuadamente la disciplina matemática con respecto a la cual debe usarse el concepto.

El caso de la explicación científica es similar, pues la investigación científica trata de explicar fenómenos empíricos por medio de leyes y teorías, que son objetivas en el sentido de que sus implicaciones empíricas y los elementos de juicio que las sustentan sean independientes de los individuos particulares que las someten a prueba o las aplican. Y las explicaciones, así como las predicciones, basadas en tales leyes y teorías pretenden ser objetivas en un sentido análogo. Este propósito ideal plantea el problema de elaborar un concepto no pragmático de explicación científica, un concepto abstraído —por así decir— del pragmático y que no exija una relativización mayor que el concepto de prueba matemática, con respecto a los individuos que indagan. Los modelos de leyes abarcantes pretenden elucidar este concepto no pragmático de la explicación.

Proponer esos modelos, no significa negar la "dimensión" pragmática de la explicación ni disminuir su importancia; tampoco, por supuesto, significa sostener que la gente hallará esclarecedora o satisfactoria una explicación sólo con que ésta se ajuste a uno de los modelos de leyes inclusivas. Para explicar un fenómeno dado a una persona, bastará a menudo llamar su atención sobre algún hecho particular del cual no ha tomado adecuado conocimiento. Esto es verdad, presumiblemente, del hombre mencionado en un periódico, hace algunos años, y quien estaba intrigado al descubrir que su casa se enfriaba cuando miraba un programa de televisión en invierno. Todo lo que se le deberá decir a título de explicación era que el equipo de televisión estaba colocado directamente debajo del termostato, y al calentarlo se cerraba la calefacción. Así, a menudo la búsqueda de una explicación es una búsqueda de la "causa" del suceso desconcertante, en el sentido vago que aquí ilustramos. El indagador que acepta como satisfactoria una explicación causal particular a veces tendrá una información básica de tipo nomológico —por ejemplo, acerca de cómo funciona un termostato— que justifique la atribución causal. En otros casos, puede carecer de tal información y, no obstante, quedar satisfecho con la explicación: las condiciones pragmáticas para la aceptabilidad de una explicación propuesta no coinciden con las condiciones lógico-sistemáticas que pretenden elucidar los modelos de leyes inclusivas. Cuando las leyes atinentes al caso son más o menos claramente comprendidas y admitidas por el indagador sería incorrecto —por supuesto— decir que su pregunta tuvo la función pragmática de producir

leyes inclusivas; pero no es incorrecto ni superfluo hacer referencia a tales leyes si debe hacerse explícita la lógica de la explicación, y especialmente la fuerza explicativa de los hechos particulares mencionados en ella.

En otros contextos —por ejemplo, con frecuencia, en la investigación científica— la preocupación pragmática que estipula la búsqueda de una explicación puede ser el deseo de descubrir leyes o principios teóricos que abarquen determinada clase de fenómenos. En otros casos, aun, el indagador puede tener conocimiento de los datos y leyes particulares requeridos, pero puede necesitar que se le muestre de qué manera es posible derivar la explicación de tal información.[108]

Pero llamar la atención hacia las importantes facetas pragmáticas de la explicación e indicar los diversos procedimientos que puedan resultar adecuados en diferentes casos para disipar la perplejidad en la búsqueda de una explicación, no equivale a mostrar que un modelo no pragmático de explicación científica debe ser irremisiblemente inadecuado, así como razonamientos análogos concernientes a la noción de prueba no muestran que los modelos no pragmáticos de prueba deben ser estériles y no esclarecedores. Como es bien sabido, sucede lo contrario.

Por lo tanto, no tiene objeto quejarse de que los modelos de leyes inclusivas no se ajusten exactamente a la forma en que los científicos presentan realmente sus explicaciones. Por lo general se eligen esas formulaciones teniendo en cuenta un auditorio determinado y, por ende, determinados requisitos pragmáticos. Esto también es verdad respecto de la manera como los matemáticos presentan sus pruebas; pero la teoría matemática de la prueba hace abstracción, justificadamente, de estas consideraciones pragmáticas.[109]

5.2 *La explicación de "cómo fue posible"*

Un importante aspecto pragmático de la explicación se refleja en la distinción de Dray entre "explicar por qué necesariamente" se produjo un suceso y "explicar cómo fue posible" que un suceso se haya producido.[110] Una explicación D-N puede considerarse adecuadamente al primer propósito; realizar lo segundo es una tarea muy diferente, como ahora veremos.

Si un amigo me dice que en una reunión a la que asistió la noche de Año Nuevo su cucharita se disolvió al colocarla en una taza de ponche caliente, puedo preguntarme: ¿cómo puede haber sucedido esto, ya que el metal no se funde a una temperatura tan baja? Análogamente, las no-

[108] En un interesante examen de aspectos de la explicación que son en gran medida pragmáticos, Scriven usa el término "explicación por derivación" para designar una explicación que consiste simplemente en demostrar esta cualidad de derivación, y da un ejemplo de la historia de la ciencia que muestra que tal derivación bien puede presentar considerables dificultades matemáticas y, por ende, ser difícil de descubrir (Scriben, 1959, págs. 461-462).

[109] Véanse también los comentarios sobre este punto en la sección 1 del artículo de Bartley (1952), en el que se defiende contra esta acusación a la exposición que hace Popper del modelo deductivo. En Pitt (1959, págs. 585-586) se hallarán observaciones en un espíritu similar.

[110] Véase Dray (1957, págs. 158 y sigs.).

ticias de que el *Andrea Doria* se había hundido como resultado de un choque planteó la cuestión de cómo pudo haber sucedido eso, considerando que el barco estaba equipado con los más modernos medios de salvamento y conducido por marinos experimentados.

Como indican estos ejemplos, normalmente nos preguntamos cómo pudo haber sucedido *X* sólo si, con palabras de Dray, "lo que sabemos parece excluir la posibilidad del suceso que debe explicarse",[111] es decir, si algunas de las creencias que abrigamos concernientes a cuestiones fácticas atinentes al caso hacen imposible, o al menos sumamente improbable en nuestro parecer, que *X* pueda haber sucedido; en esto reside el aspecto pragmático de la cuestión. Para dar una explicación satisfactoria de "cómo fue posible" un suceso, será necesario, pues, hallar las suposiciones empíricas que subyacen a la cuestión y luego mostrar que algunas de ellas son falsas o que el indagador se equivocaba al pensar que tales suposiciones eran una garantía de su creencia acerca de que *X* no pudo haber sucedido. En el caso de la cucharita bastaría señalar que algunos metales, como la aleación de Wood, se funden a la temperatura del ponche caliente; y podría lograrse una explicación completa por leyes inclusivas estableciendo que la cucharita en cuestión estaba realmente hecha con la aleación de Wood para el uso de bromistas.[112]

En el caso de *Andrea Doria,* si la pregunta "¿cómo pudo haber sucedido *X*?" surge de suposiciones que parecen hacer sumamente improbable que pueda haber sucedido *X*, pero no lo excluyen lógicamente, entonces una respuesta adecuada puede consistir en señalar que el indagador se equivoca en algunas de sus suposiciones fácticas o en la creencia de que las suposiciones hacen muy improbable que suceda *X*: estas dos posibilidades son análogas a las consideradas en el ejemplo anterior. Pero además, tenemos aquí una tercera posibilidad, sugerida también por nuestro anterior examen de la lógica de la explicación estadística: todas las suposiciones del indagador atinentes al caso pueden ser verdaderas, y su creencia de que ellas hacen muy improbable que suceda *X* puede ser correcta. En tal caso, la perplejidad expresada por la pregunta del indaga-

111 Dray (1957, pág. 161).

112 En un comentario bibliográfico sobre el libro de Dray (1958), Passmore llega a decir que "responder a una pregunta del tipo 'Cómo es posible', a menos que se la responda por una mera adivinanza, es esbozar una explicación del tipo 'por qué-necesariamente' ". Si bien esta observación parece básicamente correcta, se la debe hacer más liberal, creo, de modo que permita la elaboración de una explicación "por qué-necesariamente" o de una "por qué probablemente". Si alguien pregunta cómo puede haber sucedido *X*, por lo general no se satisfará con que se le diga simplemente que se equivocaba en algunas de sus suposiciones empíricas, de las que pensaba que excluían la producción de *X*. También querrá que se le dé un conjunto de suposiciones alternativas y presumiblemente verdaderas que, junto con el resto de sus creencias básicas, le expliquen por qué sucedió *X*. El ejemplo de la cuchara que se diluye ilustra este caso. Pero si quien nos interroga cree que volcar sal siempre va seguido de mala suerte dentro de los tres días, y si nos preguntara "¿cómo pude haber escapado a la mala suerte si volqué la sal hace tres días?", entonces la respuesta no podría ser otra que señalarle la falsedad de su hipótesis general y, quizá, que en la gran mayoría de los casos, el vuelco de la sal no va seguido por mala suerte; pero no se podrá aducir ninguna explicación del tipo "por qué necesariamente" para el hecho de que nuestro interlocutor no haya padecido de mala suerte.

dor, "¿cómo puede haber sucedido?", puede resolverse ampliando sus elementos de juicio totales, es decir, llamando su atención hacia ciertos hechos adicionales que agregados a los tomados anteriormente en consideración hacen menos improbable que suceda *X*.

Observaciones similares se aplican a las preguntas de la forma "por qué no se da el caso de que *p*?", que podrían ser reformuladas como preguntas de la forma "¿cómo fue posible?": "¿cómo fue posible el caso de que no-*p*?" Preguntas tales como "¿por qué no se viene abajo la torre inclinada de Pisa?", "¿por qué los hombres de las antípodas no se caen fuera de la tierra?" o "si la reflexión en un espejo plano cambia la izquierda por la derecha, ¿por qué no cambia también lo de arriba por lo de abajo?" normalmente sólo se plantearán si el indagador abriga ciertas suposiciones sobre cuestiones empíricas atienentes a cada caso que, según le parece, dan seguridad o, al menos, hacen muy probable que el fenómeno especificado se produzca. Nuevamente, una respuesta pragmáticamente adecuada deberá aclarar los errores lógicos o empíricos subyacentes en esta creencia.

Y, por supuesto, las preguntas que piden una explicación y forma típica "¿por qué se da el caso de que *p*?", a menudo —aunque no invariablemente— obedecen a la creencia de que no sería el caso de que *p*, creencia que, nuevamente, puede parecer al indagador más o menos firmemente apoyada por otras suposiciones empíricas que acepta como verdaderas. Y en esta eventualidad, el indagador puede no sentirse satisfecho si se le ofrece simplemente, por ejemplo, una explicación por leyes inclusivas de por qué se da el caso de que *p*. Para disipar su perplejidad quizá sea menester mostrarle que algunas de las suposiciones subyacentes en su expectación contraria eran erróneas.[113]

5.3 *La explicación frente a reducción a lo familiar*

Una concepción predominantemente pragmática de la explicación como tendiente a disipar el desconcierto del indagador, subyace también a la difundida idea de que una explicación debe, de alguna manera, reducir o vincular el fenómeno desconcertante a algo que ya es familiar para el indagador y que acepta como no problemático. Así, por ejemplo, Bridgman sostiene que "la esencia de una explicación consiste en reducir una situación a elementos con los que estamos tan familiarizados que los aceptamos como cosa obvia, de modo que nuestra curiosidad se satisfaga".[114]

Un examen de esta caracterización explícitamente pragmática puede contribuir también a aclarar y dar apoyo a las razones para construir un concepto no pragmático de explicación científica.

113 Este aspecto de la explicación, y otros relacionados con él, han sido aguda y claramente examinados por S. Bromberg (1960). Sugerentes observaciones sobre los aspectos pragmáticos de la explicación se hallarán también en Passmore (1962).

114 Bridgman (1927, pág. 37). El carácter pragmático de esta concepción se refleja claramente en la observación de Bridgman de que "una explicación no es una cosa absoluta, sino que lo que es satisfactorio para una persona no lo es para otra". *Loc. cit.*, pág. 38.

Es innegable que muchas explicaciones científicas realizan, en cierto sentido, una "reducción a lo familiar". Podría decirse esto, por ejemplo, de la explicación teórico-ondulatoria de la refracción y la interferencia ópticas, y al menos de algunas de las explicaciones logradas por la teoría cinética del calor. En los casos de este tipo, los conceptos y principios invocados en el explanans presentan una semejanza más o menos estrecha con conceptos y principios que se han usado desde hace tiempo en la descripción y explicación de algún tipo familiar de fenómeno, tal como la propagación del movimiento ondulatorio sobre la superficie del agua o el movimiento de bolas de billar.

En lo concerniente a la concepción general de la explicación como una reducción a lo familiar, observemos ante todo que lo que es familiar para una persona puede no serlo para otra y, por lo tanto, que este enfoque conciba la explicación como algo relativo a un indagador. Pero, como señalamos antes, las explicaciones del tipo buscado por la ciencia empírica pretenden poner de manifiesto relaciones objetivas.

En segundo lugar, la concepción que estamos examinando sugiere que lo familiar no requiere explicación alguna. Pero esta idea no está de acuerdo con el hecho de que los científicos hayan dedicado grandes esfuerzos a explicar fenómenos "familiares", tales como los cambios de las mareas, el rayo, el trueno, la lluvia, la nieve, el color azul del cielo, las semejanzas entre los padres y su progenie, el hecho de que la luna parezca mucho más grande cuando está cerca del horizonte que cuando está alto en el cielo, el hecho de que ciertas enfermedades sean "contagiosas" y otras no, y hasta el hecho muy familiar de que de noche esté oscuro. En verdad, la oscuridad del cielo nocturno parece un fenómeno que necesita mucho de una explicación, en vista de la paradoja de Olbers. Este razonamiento, expuesto en 1826 por el astrónomo alemán Heinrich Olbers, se basa en unas pocas suposiciones simples que son, en términos aproximados: las distancias y las luminosidades intrínsecas de las estrellas presentan aproximadamente la misma distribución de frecuencia en el pasado que en el presente, a través de todo el universo; la leyes básicas de la propagación de la luz se cumplen en todas las regiones espacio-temporales del universo; el universo en su conjunto es estático, esto es, no se producen en él movimientos sistemáticos en gran escala. A partir de estas suposiciones se sigue que el cielo en todas direcciones y en todos los tiempos debe tener un enorme brillo uniforme, y que la energía que llega a la superficie de la Tierra debe corresponder a una temperatura de más de 5500°C.[115]

Por consiguiente, la paradoja de Olbers plantea una pregunta del tipo "¿cómo es posible que...?" Una respuesta sugiere la reciente teoría de que el universo se expande constantemente. Esta teoría implica, primero, que la suposición de Olbers referente al carácter estático del universo es errónea, y suministra, en segundo lugar, una explicación positiva de la oscuridad del cielo nocturno, al mostrar que la energía de la radia-

[115] Una exposición más completa de la paradoja y un análisis crítico realizado a la luz de las actuales teorías cosmológicas se hallarán, por ejemplo, en Bondi (1961, capítulo 2) y Sciama (1961, capítulo 6).

ción que llega de estrellas muy distantes se reduce enormemente por causa de las grandes velocidades a las que se alejan.

Este ejemplo también aclara otro punto, a saber, que en lugar de reducir lo no familiar a lo familiar, una explicación cintífica a menudo hace lo opuesto: explica fenómenos familiares con ayuda de concepciones teóricas que pueden parecer poco familiares y hasta contrarias a la intuición, pero que dan cuenta de una gran variedad de hechos y reciben adecuado apayo de los resultados de los tests científicos.[116]

Estas observaciones son aplicables también fuera del dominio de las ciencias naturales. Su importancia para la sociología, por ejemplo, se sugiere en el pasaje inicial de un libro de Homans: "Mi tema es un caos familiar. Nada es más familiar para los hombres que su conducta social ordinaria y cotidiana... cada hombre hace sus propias generalizaciones acerca de su experiencia social, pero las usa *ad hoc* dentro del ámbito de situaciones al cual se aplica cada una de ellas, las abandona tan pronto como termina su valor inmediato y nunca se pregunta cómo se relacionan entre sí... El propósito de este libro es extraer del caos familiar algún orden intelectual".[117] Dicho sea de paso, Homans continúa diciendo que el ordenamiento requerido de un conjunto de hechos sociológicos empíricamente establecidos, representado por generalizaciones de bajo nivel, exige un *explicación* de esos hechos; y que logra tal explicación mediante un "conjunto de oraciones más generales, aunque de la misma forma que las empíricas, a partir de las cuales sea posible deducir lógicamente éstas en condiciones específicas determinadas. Deducirlas exitosamente es explicarlas".[118]

A este énfasis en el interés del sociólogo por la explicación teórica de generalizaciones "familiares" acerca de la conducta social, debe agregarse una observación que ha sido destacada por Lazarsfeld, entre otros; a saber, que lo que muchos consideran como hechos obvios y familiares de la experiencia psicológica y sociológica cotidiana, a veces no son hechos en absoluto, sino estereotipos populares. Esto es verdad, para no mencionar más que uno de los interesantes ejemplos de Lazarsfeld, de la idea según la cual el intelectual es emocionantemente menos estable que el hombre de la calle, psicológicamente más impasible, y que era de esperar, por ende, entre los soldados de los Estados Unidos de la Segunda Guerra Mundial, que los hombres de mayor educación manifestaran más síntomas psiconeuróticos que los menos educados. De hecho, se halló que sucede lo opuesto.[119] Así, una explicación de un caso particular basada en la generalización de bajo nivel de este estereotipo es simplemente falsa, aunque pueda decirse que efectúa una reducción a lo familiar.

Tal reducción, pues, según se desprende de nuestra extensa argumentación, no es sin duda una condición necesaria para que una explicación

[116] También se destaca este punto en el conciso y aclarador artículo de Feigl (949); y se lo ilustra lúcidamente con referencia a la teoría de la relatividad en Frank (1957, págs. 133-134).
[117] Homans (1961, págs. 1-2).
[118] Homans (1961, págs. 9-10); bastardillas del autor.
[119] Véase Lazarsfeld (1949, págs. 379-380).

científica sea aceptable. Pero tampoco es una condición suficiente, pues a veces puede responderse al pedido de una explicación de una manera que satisfaga la curiosidad del interrogador provocando en él una sensación de familiaridad con respecto a un fenómeno inicialmente desconcertante, sin brindarle una explicación científicamente aceptable. En este caso, podría decirse que la familiaridad ofrece satisfacción, pero no comprensión. Por ejemplo, como acabamos de ver, la explicación expresada podría basarse en una creencia familiar pero equivocada, y por ende será falsa. O bien la explicación propuesta podría basarse en ideas metafóricas y metafísicas inestables, y no en hipótesis empíricas generales, por lo cual ni siquiera brindaría una potencial explicación científica. Tomemos como ejemplo la "hipótesis de un subconsciente común", propuesta para explicar presuntos fenómenos telepáticos.[120] Ella afirma que si bien en el dominio consciente las mentes humanas son entidades separadas, están vinculadas por un subconsciente común, del cual emergen las conciencias individuales como islotes montañosos unidos por un continente submarino. El sugestivo conjunto de imágenes de esta explicación puede despertar la sensación de una comprensión intuitiva de los fenómenos telepáticos; éstos parecen explicados por reducción a ideas con las que estamos muy familiarizados. Sin embargo, lo que se nos ha dado es un símil, no una explicación científica. La explicación propuesta no nos ofrece ningún fundamento para esperar razonablemente la aparición de fenómenos telepáticos, ni nos brinda clave alguna acerca de las condiciones en las cuales es probable que aparezcan tales fenómenos. En verdad, en la forma aquí esboraza, la noción de subconsciente común no tiene implicaciones claras concernientes a fenómenos empíricos y, por lo tanto, no se la puede someter a tests objetivos o a un uso explicativo y predictivo que tenga significación.

Una crítica similar puede hacerse de las explicaciones neovitalistas de ciertos fenómenos biológicos en términos de entelequias o fuerzas vitales. Tales explicaciones no especifican en qué condiciones una fuerza vital ejercerá su influencia ni qué forma específica adoptarán sus manifestaciones; y en el caso de interferencia externa en un organismo, tampoco determinan en qué medida una entelequia compensará la perturbación resultante. Por el contrario, una explicación de los movimientos planetarios en función de la teoría newtoniana de la gravedad especifica qué fuerzas gravitacionales ejercerán sobre un planeta determinado el Sol y otros planetas, conocidas sus masas y sus distancias, y precisa además qué cambios en el movimiento cabe esperar como resultado de la acción de esas fuerzas. Ambas explicaciones aducen ciertas "fuerzas" que no es posible observar directamente: una de ellas, fuerzas vitales; la otra, fuerzas gravitacionales; sin embargo, la segunda tiene status explicativo, mientras que la primera no lo tiene. Esto es un consecuencia de hecho de que la teoría newtoniana brinda leyes específicas que gobiernan las fuerzas gravitacionales, mientras que el neovitalismo no especifica ninguna

[120] Véase la referencia crítica en Price (1945) y el uso que hace Carington de la idea como "una analogía" (1949, pág. 223 y sigs.), así como su explicación más específica del concepto de subconsciente común (*loc. cit.*, págs. 208 y sigs.).

ley que gobierne las fuerzas vitales: sólo es metafórico, en efecto. Por consiguiente, lo fundamental para una explicación científica son las leyes inclusivas o los principios teóricos, no la sensación de familiaridad que puedan transmitir sus palabras.

Las leyes invocadas en una explicación científica propuesta son, por supuesto, testables; y los resultados adversos pueden conducir a su rechazo. Ningún destino semejante amenaza a las explicaciones dadas en términos de analogías o metáforas: puesto que no especifican qué es lo que cabe esperar en condiciones empíricas, ningún test empírico puede desacreditarlas. Pero la absoluta inmunidad a la desconfirmación no es una ventaja sino un derecho fatal, cuando lo que nos interesa es llegar a un cuerpo de conocimiento empírico objetivamente testable y empíricamente bien sustentado, como en la investigación científica. Una explicación que no tiene implicaciones concernientes a fenómenos empíricos no puede servir a este propósito, por fuerte que sea su llamado a la intuición: desde el punto de vista de la ciencia es una *seudo-explicación*, una explicación sólo aparente.

En resumen, pues, no es necesario ni suficiente, para que una explicación sea científicamente adecuada, que reduzca el explanandum a ideas con las que ya estamos familiarizados.

6. *Modelos y analogías en la explicación científica*

Las explicaciones que brindan las ciencias empíricas a veces están formuladas en función de un "modelo" de los fenómenos que se quiere explicar, o en función de analogías entre estos fenómenos y otros anteriormente estudiados. En esta sección me propongo examinar algunas formas de este procedimiento y estimar su valor explicativo.

Consideremos en primer término el uso —muy difundido en el siglo XIX y a comienzos del XX— de sistemas mecánicos más o menos complejos como modelos de fenómenos eléctricos, magnéticos y ópticos, del éter luminífero, etc. La importancia que algunos eminentes científicos atribuyeron a tales representaciones se refleja en la famosa declaración de sir William Thomson (luego Lord Kelvin):

> Nunca me siento satisfecho hasta que puedo hallar un modelo mecánico de algo. Si puedo elaborar un modelo mecánico, puedo entenderlo. En la medida en que no puedo elaborar un modelo mecánico a lo largo de todo el proceso, no puedo comprenderlo...[121]
>
> Mi propósito es mostrar cómo elaborar un modelo mecánico que cumpla con las condiciones exigidas en los fenómenos físicos que estudiamos, sean cuales fueren. Cuando estudiemos el fenómeno de la elasticidad en los sólidos, mostraré un modelo de éste. En otro momento, cuando estudiemos las vibraciones de la luz, mostraré un modelo de la acción que se manifiesta en este fenómeno... Creo que

[121] Thomson (1884, págs. 270-271).

el test de: "¿Comprendemos o no un punto particular de la física?" es "podemos elaborar un modelo mecánico de él?" [122]

Sir Oliver Lodge, cuyo libro sobre la electricidad presenta una multitud de modelos mecánicos, dice en un espíritu similar:

> Pensad en los fenómenos eléctricos como si estuvieran producidos por un líquido que todo lo impregna introducido en una gelatina; en una máquina eléctrica como una bomba, en la carga como un exceso o un defecto, en la atracción como debida a una tensión, en la descarga como una explosión... Pensando de este modo, obtendréis una comprensión más real del tema y una mayor visión de los procesos que se producen en la naturaleza —por desconocidos que puedan ser éstos, hablando estrictamente— que si emplearais las viejas ideas de acción a distancia, o que si os contentarais con no poseer ninguna teoría para vincular los hechos... Estoy convencido de que no es razonable ir a la deriva entre una multitud de fenómenos complicados sin más guía que la suministrada por áridas y rígidas ecuaciones matemáticas.[123]

Estas declaraciones son variantes de la idea de que la explicación, en la ciencia, debe implicar una reducción a lo familiar. Lo que estas variantes particulares exigen no es simplemente que una explicación haga, de alguna manera, plausible o familiar un fenómeno, sino más específicamente que suministre un modelo gobernado por las leyes de la mecánica, a las que en este contexto se les otorga el status de principios familiares.

Pero, ¿qué es exactamente lo que consigue la construcción de un modelo mecánico? Por supuesto, no se pretende identificar el fenómeno modelado con el modelo. No se sostiene que una corriente eléctrica mantenida en un alambre por medio de una batería sea lo mismo que el fluir de un líquido a través de tubos mantenido por medio de una bomba, o que una cuerda inextensible que se mantiene en movimiento por medio de poleas a las que mueve un peso sumergido.[124] Lo que se afirma simplemente es que se obtiene una analogía entre el modelo y el fenómeno que representa. Y el aspecto importante de la analogía reside en una semejanza formar entre ciertas leyes que gobiernan el sistema mecánico y las leyes correspondientes del fenómeno modelado.

Consideremos, por ejemplo, la analogía citada a menudo entre el flujo de una corriente eléctrica por un alambre y el flujo de un líquido por un tubo. Si el líquido fluye con moderada velocidad a través de un tubo bastante estrecho de sección circular, entonces, según la ley de Poiseulle, el volumen V del líquido que fluye por segundo por una sección determinada es proporcional a la diferencia de presión entre los extremos del tubo:

[122] Thomson (1884, págs. 131-132).
[123] Lodge (1889, págs. 60-61).
[124] Puede hallarse una gran profusión de tales modelos en Lodge (1889), y en Thomson (1884). S. B. Russell (1913) describe un modelo hidrodinámico que representa de una manera muy similar ciertos aspectos de la conducta del sistema nervioso.

(6.1a) $V = c \cdot (p_1 - p_2)$.

Esta ley tiene la misma forma que la ley de Ohm sobre el flujo de electricidad por un conductor metálico:

(6.1b) $I = k \cdot (v_1 - v_2)$.

Puede decirse que en esta última, la fuerza de la corriente, I, representa la cantidad por segundo de carga eléctrica que fluye por una sección determinada del alambre; $v_1 - v_2$ es la diferencia de potencial entre los extremos del alambre; y k es el recíproco de la resistencia.

La analogía va más allá aún. El factor c de (6.1a) es inversamente proporcional a la longitud l_1 del tubo:

(6.2a) $$c = \frac{c'}{l_1}$$

y análogamente, el factor k de (6.1b) es inversamente proporcional a la longitud, l_2, del alambre:

(6.2b) $$k = \frac{k'}{l_2}$$

Así, la analogía en virtud de la cual el flujo de un líquido constituye un modelo del flujo de una corriente puede caracterizarse del siguiente momo: un cierto conjunto de leyes que gobiernan el primer fenómeno tiene la misma estructura sintáctica que un conjunto correspondiente de leyes del segundo fenómeno; o más explícitamente, los términos empíricos (es decir, aquellos que no son lógicos ni matemáticos) [125] que aparecen en el primer conjunto de leyes pueden aparearse, uno por uno, con los del segundo conjunto, de tal modo que si en una de las leyes del primer conjunto cada término reemplaza a su análogo, se obtiene una ley del segundo conjunto, y viceversa. Diremos que dos conjuntos de leyes de este tipo son sintácticamente isomórficas. Dicho brevemente, pues, la semejanza o "analogía" importante entre un modelo del tipo aquí considerado y el tipo de fenómeno modelado consiste en un *isomorfismo nómico*, es decir, un *isomorfismo sintáctico entre dos conjuntos correspondientes de leyes*. La noción de modelo así obtenida no se limita a los sistemas mecánicos, por supuesto; en el mismo sentido, podemos hablar también de modelos eléctricos, químicos y aun de otros tipos de "modelos analógicos".

Pero en nuestro ejemplo, como en otros casos de modelos analógicos, el isomorfismo tiene sus límites: algunas de las leyes para el flujo de un líquido por tubos no son válidas para las corrientes eléctricas que fluyen por alambres. Por ejemplo, si la longitud del tubo y la diferencia de presión entre sus extremos son fijas, V es proporcional a la cuarta potencia del radio de la sección transversal, mientras que en circunstancias seme-

[125] Las constantes físicas tales como "s" y "q" en (6.3b) son aquí términos empíricos.

jantes la corriente es proporcional al cuadrado de la sección transversal del alambre:

$$(6.3a) \quad V = \frac{\pi r_1^4}{8 l_1 s} (p_1 - p_2)$$

$$(6.3b) \quad I = \frac{\pi r_2^2}{l_2 q} (v_1 - v_2)$$

Aquí, s es la viscosidad del líquido y q la resistencia específica del metal del que está hecho el alambre; r_1 es el radio de la sección transversal del tubo, y r_2 es el radio del alambre.

Así, la afirmación de que un sistema S_1 es un modelo analógico de un sistema S_2 es elíptica. Una oración completa que exprese esa relación debería adoptar la forma: "S_1 es un modelo analógico de S_2 con respecto a los conjuntos de leyes L_1 y L_2". Esta oración es verdadera si las leyes de L_1 se aplican a S_1 y las de L_2 a S_2 y si L_1 y L_2 son sintácticamente isomorfas.[126]

El concepto de analogía como un isomorfismo nómico desempeña un papel importante en el ensayo de Maxwell sobre las líneas de fuerza Faraday. En él Maxwell dice: "Entiendo por analogía física una semejanza parcial entre las leyes de una ciencia y las de otra en virtud de la cual cada una de ellas es un ejemplo de la otra". En lo concerniente a la analogía entre la luz y las vibraciones de un medio elástico, observa que "si bien no puede sobreestimarse su importancia y su fecundidad, debemos recordar que sólo se funda sobre una semejanza *de forma* entre las leyes de la luz y las de las vibraciones".[127] Y continúa: "Mediante el uso de analogías de este tipo he tratado de exponer, en una forma conveniente y práctica, las ideas matemáticas que son necesarias para el estudio de los fenómenos de la electricidad... No pretendo fundar una teoría física..., y... el límite de mi propósito es mostrar cómo por una estricta aplicación de las ideas y métodos de Faraday, puede exponerse con claridad, ante la mente matemática, la conexión entre los muy diferentes órdenes de fenómenos que él descubrió".[128] La analogía que luego Maxwell desarrolla en detalle se basa en una representación de las líneas de fuerza de Faraday mediante tubos a través de los cuales fluye un líquido incompresible. Es interesante observar que si bien Maxwell puede ofrecer una representación analógica de muchos fenómenos eléctricos y mag-

[126] Esta caracterización de los modelos analógicos concuerda con las concepciones de Maxwell y Duhem sobre la analogía en física, acerca de las cuales volveremos a referirnos en seguida. También recibe sustento de la manera como Boltzmann (1891) utiliza los modelos mecánicos para representar el ciclo de Carnot en la teoría del calor (1891, capítulo 2) y diversos fenómenos eléctricos. El concepto general de "modelo dinámico" de Heinrich Hertz refleja la misma idea básica; véase Hertz (1894, pág. 197).

[127] Maxwell (1864, pág. 28); bastardillas del autor.

[128] Maxwell (1864, pág. 29).

néticos, se halla en la imposibilidad de extender la analogía cuando llega al examen de lo que Faraday había llamado el estado electrotónico; en este punto, recurre a la formulación de una teoría en términos puramente matemáticos.[129]

Las ideas de hombres como Kelvin y Lodge en lo concerniente a la importancia de los modelos analógicos para la explicación en física fueron severamente criticadas por Duhem. Este ve el objetivo de la física en la construcción de teorías formuladas en términos matemáticos precisos de las cuales puedan deducirse leyes establecidas empíricamente, y arguye que los modelos matemáticos no contribuyen en nada al logro de este objetivo. Con referencia al libro de Lodge, Duhem comenta: "He aquí un libro que pretende exponer las teorías modernas de la electricidad...; sólo nos habla de cuerdas que se mueven por poleas, que se enrollan en tambores, que atraviesan bolillas y transportan pesos; de tubos que bombean agua y de otros que se dilatan y se contraen; de ruedas dentadas que se engranan e impulsan cremalleras. Creíamos haber entrado en la pacífica y bien ordenada mansión de la razón, y nos encontramos en una fábrica".[130] Duhem continúa quejándose de que lejos de facilitar la comprensión de una teoría, "para un lector francés", el uso de tales modelos mecánicos le exige un serio esfuerzo para comprender el funcionamiento del complicado aparato y reconocer las analogías entre las propiedades del modelo y la teoría que se trata de aclarar.

Aunque Duhem rechaza el uso explicativo de los modelos mecánicos, subraya que las analogías, en cambio, pueden resultar muy fructíferas en la investigación física. Las analogías a las que alude son las que se basan en lo que hemos llamado isomorfismo nómico. Menciona, por ejemplo, la transferencia realizada por Ohm de las leyes de la conducción de calor a la conducción eléctrica, y destaca la importancia de esos casos en los que vastas teorías referentes a dos categorías distintas y diferentes de fenómenos tienen la misma forma algebraica.[131]

Sin embargo, si nuestra caracterización es correcta, los modelos mecánicos de los que Duhem se burla manifiestan isomorfismos nómicos que son, básicamente, del mismo tipo que las analogías científicas de Duhem, no formuladas específicamente en el lenguaje de los modelos. La distinción de Duhem entre modelos y analogías, distinción que no basa en criterios precisos, no refleja, pues, una diferencia de status lógico, sino más bien una diferencia en la precisión y el alcance de los conjuntos isomórficos de leyes. Entre las leyes que gobiernan un modelo mecánico, aquellas que pueden trasladarse isomórficamente al fenómeno modelado por lo común son pocas y de alcance limitado, de modo que a veces se

[129] Maxwell (1864, págs. 51 y sigs.). Para un examen más completo de las ideas de Maxwell sobre la importancia de las analogías para la teoría física, véase los estudios de Turner (1955 y 1956).

[130] Traducido de Duhem (1906, pág. 111).

[131] Duhem (1906, págs. 152-154). Boltzmann caracteriza las analogías físicas de una manera similar: "...La naturaleza, por así decir, ha construido las cosas más diversas siguiendo exactamente el mismo plan; o como dice secamente el matemático analista, las mismas ecuaciones diferenciales son válidas para los más diversos fenómenos". Traducido de Boltzmann (1905, pág. 7).

utilizan varios modelos diferentes para representar distintos aspectos de un mismo tipo de entidad o fenómeno físico. Por ejemplo, Kelvin ofrece modelos mecánicos de las moléculas muy diferentes para representar la elasticidad en los cristales, la dispersión de la luz y la rotación del plano de polarización de un haz luminoso.[132] Y Lodge diseña modelos mecánicos muy diferentes, del tipo al que alude Duhem en el pasaje citado, para representar diversos fenómenos electrostáticos, electrodinámicos y electromagnéticos. En el caso de analogía fecundas del género considerado por Duhem, en cambio, las leyes o principios teóricos isomórficos son enunciados en términos matemáticos precisos y tienen fuerza suficiente como para permitir la deducción de un gran número de consecuencias variadas que constituyen, en sí mismas, leyes importantes. Ejemplifican esto los extensos isomorfismos nómicos que permiten la aplicación de la teoría matemática de los movimientos ondulatorios a ciertas partes de la mecánica, la óptica y la mecánica cuántica.[133]

Para estimar la importancia explicativa de los modelos analógicos, y con mayor generalidad de las analogías basadas en isomorfismos nómicos, supongamos que se está explorando algún campo de investigación "nuevo" y que tratamos de explicar los fenómenos que hallamos en él mediante una referencia analógica a algún dominio de la investigación "viejo" y ya explorado. Esto exige la determinación de un isomorfismo entre un conjunto de leyes, L_1, pertenecientes al viejo campo y un conjunto correspondiente, L_2, del nuevo. Para ello, obviamente debemos primero descubrir un conjunto L_2 adecuado de leyes del nuevo campo. Pero una vez hecho esto, esas leyes pueden usarse directamente para la explicación de los "nuevos" fenómenos, sin ninguna referencia a su isomorfismo estructural con el conjunto L_1. Para los propósitos sistemáticos de la explicación científica, no es esencial recurrir a analogías y, por ende, puede prescindirse de ellas.

Esta observación se aplica también a los modelos analógicos de tipo no mecánico, como los sistemas fisicoquímicos que se han utilizado para imitar fenómenos que se consideran a menudo como específicamente biológicos. Leduc, por ejemplo,[134] logró producir por medios puramente químicos una gran variedad de crecimientos osmóticos cuyas formas muy diversificadas se asemejan de manera sorprendente a los de plantas y animales que nos son familiares, y que en su desarrollo manifiestan notables analogías con los crecimientos orgánicos. Los modelos analógicos así obtenidos se basan en un isomorfismo de leyes no cuantitativas:

> Un crecimiento osmótico tiene una existencia evolutiva. Se nutre por ósmosis e intususcepción, realiza una elección entre las sustancias que se le ofrecen y modifica la constitución química de su alimento antes de asimilarlo. Al igual que un ser vivo, arroja a su

[132] Véase Thomson (1884).
[133] Otros ejemplos de analogías basadas en isomorfismos nómicos en la física se encontrarán en el artículo de Seeliger (1948). Para un esclarecedor examen, bien ilustrado, de la significación del isomorfismo nómico en la física, véase también Watkins (1938, capítulo 3).
[134] Véase los libros profusamente ilustrados de Leduc (1911 y 1912).

ambiente los residuos de su función. Además, crece y desarrolla estructuras como las de los organismos vivos, y es sensible a muchos cambios exteriores que influyen sobre su forma y su desarrollo. Pero son justamente estos fenómenos —la nutrición, la asimilación, la sensibilidad, el crecimiento y la organización— los que se consideran generalmente como las únicas características de la vida.[135]

Estas analogías, y otras más, entre organismos y sistemas fisicoquímicos han sido usadas a menudo para responder a la afirmación vitalista de que el crecimiento, el metabolismo, la regeneración, etc., son fenómenos que no manifiestan una "máquina" o un sistema gobernado exclusivamente por leyes fisicoquímicas.[136] Pero si bien los modelos pueden refutar tal afirmación, ellos no brindan una explicación teórica positiva de los fenómenos biológicos en cuestión. De hecho, Leduc ni siquiera enuncia leyes fisicoquímicas que expliquen las sorprendentes formas vegetales que presentan algunos de los crecimientos osmóticos que él provoca por medios químicos; menos aún demuestra, por supuesto, que las mismas leyes expliquen también las formas de las plantas "naturales" de los que son modelos esos crecimientos artificiales. Comentarios similares pueden hacerse del "metabolismo", la "regeneración", etc., en crecimientos osmóticos y orgánicos.

Además, los isomorfismos que presentan los modelos de Leduc y otros similares sólo conciernen a regularidades de un vago carácter cualitativo ejemplificadas por el pasaje citado: los organismos crecen y decaen, y lo mismo sus análogos osmóticos; hay un intercambio de materiales entre el organismo y el ambiente, y entre cada modelo y su ambiente; tanto en los organismos como en sus modelos fisicoquímicos se reparan en cierta medida los daños, etc. A causa de su falta de especificidad, las generalizaciones de este tipo no tienen mucho fuerza explicativa. A este respecto, las analogías expuestas son muy inferiores a las que existen entre las ondas de agua y las ondas electromagnéticas, por ejemplo, que se basan en un isomorfismo sintáctico de dos vastas teorías formuladas en términos matemáticos.

Como señalamos antes, es posible prescindir de toda referencia a analogías o modelos analógicos en la enunciación sistemática de explicaciones científicas. Pero el descubrimiento de un isomorfismo entre diferentes conjuntos de leyes o principios teóricos puede resultar útil en otros aspectos.

En primer lugar, puede contribuir a la "economía intelectual": [137]

[135] Leduc (1911, pág. 159).

[136] Véase, por ejemplo, la analogía con el cristal, examinada en Bertalanffy (1933, págs. 100-102); y también el instructivo análisis de los modelos fisicoquímicos de los fenómenos biológicos en Bonheoffer (1948), donde se indica explícitamente la motivación mencionada. En este contexto, podríamos mencionar también algunos modelos físicos de ciertos aspectos del aprendizaje, cuya construcción, también, se halla estimulada, al menos en parte, por el deseo de responder a las tesis vitalistas y otras semejantes; se encontrarán tales modelos en Baernstein y Hull (1931), y en Krueger y Hull (1931).

[137] Duhem (1906, pág. 154).

si ciertas leyes que gobiernan una "nueva" clase de fenómenos son isomórficas con las de otra clase, que ya han sido estudiadas en detalle, entonces todas las consecuencias lógicas de éstas pueden transferirse al nuevo dominio simplemente reemplazando todos los términos no lógicos por sus análogos. Un importante estudio de Gauss [138] toma como punto de partida la observación de que las fuerzas gravitacionales de atracción y las de atracción y repulsión eléctrica y magnética entre dos "elementos" son todas inversamente proporcionales al cuadrado de sus distancias y directamente proporcionales al producto de sus masas, o cargas eléctricas, o fuerzas magnéticas, respectivamente. Sobre la base de este isomorfismo nómico, Gauss desarrolla una teoría matemática general para todas las fuerzas regidas por una ley de la forma especificada, y en particular para los potenciales correspondientes, sin establecer distinción alguna entre los diferentes ámbitos a los que la teoría resultante pueda aplicarse.[139] Este aspecto del isomorfismo nómico ha hallado recientemente importantes aplicaciones prácticas en la construcción de computadoras analógicas y mecanismos similares. Por ejemplo, el isomorfismo que subyace a la analogía entre el flujo de un líquido por un tubo y el flujo de una corriente eléctrica por un alambre permite al diseñador de un vasto y costoso sistema de bombeo de agua determinar las características óptimas de las bombas y la red de tubos mediante pequeños y baratos análogos eléctricos.

Las analogías y los modelos basados en isomorfismos nómicos pueden también facilitar la comprensión de un conjunto de leyes o principios teóricos explicativos para un nuevo dominio de investigación, poniendo de manifiesto un paralelismo con principios explicativos de un dominio más familiar: de esta manera, pueden contribuir a la efectividad pragmática de una explicación.

Lo más importante aún es que las analogías o modelos bien elegidos pueden resultar útiles "en el contexto del descubrimiento", esto es, pueden brindar una guía heurística efectiva en la búsqueda de nuevos principios explicativos. Así, aunque un modelo analógico en sí mismo no explica nada, puede sugerir extensiones de la analogía sobre la cual se basó originalmente. Norbert Wiener menciona un caso de este tipo. Una analogía en la que habían pensado él y Bigelow, entre ciertos tipos de conducta humana voluntaria y la conducta de una máquina gobernada por un sistema de realimentación negativa, les sugirió la posibilidad de que existieran, para la conducta internacional, condiciones análogas a las que se dan en un sistema de realimentación que efectúa una serie de oscilaciones incontroladas, condiciones éstas que se conocen bien, teóricamente. Tal analogía se encontró, en efecto, en el estado patológico del temblor intencional, en el cual un paciente que trata de asir un objeto yerra y luego

138 Gauss (1840).

139 El descubrimiento y la utilización de isomorfismos nómicos entre campos diferentes de la investigación es uno de los objetivos de la "teoría general de sistemas", tal como la concibió Bertalanffy; véase sus breves declaraciones (1951 y 1956), donde se encontrarán referencias adicionales. En Hempel (1951a) hay algunos comentarios sobre el programa de explorar isomorfismos de la manera contemplada por Bertalanffy.

comienza a tener oscilaciones incontrolables.[140] Para dar otro ejemplo, Maxwell parece haber llegado a sus ecuaciones del campo electromagnético por el juicioso uso de analogías mecánicas de fenómenos electromagnéticos. Esto llevó a Boltzmann a decir que la alta estima que tenía Heinrich Hertz por las realizaciones teóricas de Maxwell se debió principalmente al ingenio de éste para concebir analogías mecánicas fecundas más que a su análisis matemático.[141]

Las analogías pueden resultar útiles para elaborar y desarrollar teorías de microestructuras, como la teoría cinética del calor o la teoría que explica la codificación y transmisión de información genética en función de hipótesis específicas acerca de la estructura molecular de los genes. Debe observarse, sin embargo, que tales teorías pretenden explicar uniformidades macrofísicas observables mediante adecuadas suposiciones acerca de las estructuras y procesos microfísicos subyacentes y que éstas no se presentan, por norma, solamente como modelos analógicos. Cuando Lord Kelvin trató de explicar las uniformidades en la absorción y dispersión de la luz concibiendo cada una de las moléculas que intervienen en estos procesos según el modelo de un conjunto de esferas metálicas rígidas encajonadas una dentro de otras y separadas entre sí por resortes, no pretendía por supuesto, describir la microestructura real de la materia, y habría sido ajeno a la cuestión pedirle elementos de juicio en apoyo de la suposición de que las moléculas consisten en esferas metálicas encajadas unas dentro de otras y en resortes. Sin embargo, la teoría cinética de los gases afirma, entre otras cosas, que un gas consiste en moléculas en movimiento rápido; especifica las cantidades y las masas de las partículas que intervienen, la distribución de sus velocidades y su dependencia de la temperatura, los caminos libres medios de las moléculas y el intervalo de tiempo medio entre choques sucesivos, etc.; y con respecto a éstas y muchas otras implicaciones específicas, tiene sentido pedir elementos de juicio y, en verdad, es posible suministrarlos.

Análogamente, las teorías acerca de las partículas elementales que constituyen los núcleos atómicos de diversos elementos o acerca de la estructura molecular de los genes son presentadas como explicaciones de la estructura real de los sistemas aludidos, y no como modelos analógicos. Como cualquier otra teoría de las ciencias empíricas, se proponen tales teorías de las microestructuras "hasta nuevo aviso", es decir, con la reserva de que deberán modificarse o abandonarse totalmente a la luz de posteriores elementos de juicio desfavorables que se descubran; y a menudo sólo se las presenta como aproximaciones. Sin embargo, difieren en el aspecto indicado de las explicaciones formuladas en función de modelos analógicos.

En algunas teorías de microestructuras, se supone que los componentes básicos de los macrofenómenos estudiados están gobernados por leyes idénticas o sintácticas isomórficas con respecto a un conjunto de leyes

[140] Véase Wiener (1948, págs. 13-15 y el capítulo 4).

[141] Boltzmann (1905, pág. 8; y también 1891, pág. iii). Otros ejemplos y un claro examen general del papel de las analogías en la teorización física se hallarán en Nagel (1961, págs. 107-117).

que gobiernan un campo ya explorado de la investigación. Un ejemplo característico es la suposición de que los movimientos y choques de las moléculas de un gas cumplen las leyes para los movimientos y choques de bolas de billar elásticas. Algunos autores hasta han insistido en que las suposiciones o ecuaciones básicas de toda buena teoría científica deben presentar este tipo de analogías. Un elocuente defensor de esta tesis es el físico N. R. Campbell.

Campbell considera que la principal función de las teorías es brindar explicaciones deductivas de leyes, es decir, de "proposiciones que afirman uniformidades descubiertas por experimento o por observación.[142] Una teoría, para él, consiste en dos conjuntos de proposiciones, a los que llama las hipótesis y el diccionario. La hipótesis se formula en términos de "ideas que son características de la teoría", o en términos de conceptos teóricos, como podríamos decir. El diccionario suministra una interpretación física de las hipótesis traduciendo algunas de sus proposiciones —no necesariamente todas— a otras que no suponen conceptos teóricos y que pueden certificarse o refutarse sin referencia alguna a la teoría, mediante adecuados experimentos u observaciones.[143]

Campbell exige de una teoría científica que ésta pueda explicar leyes establecidas empíricamente: tal explicación consiste en deducir las leyes a partir de las hipótesis y en conjunción con el diccionario. "Pero —sostiene— para que una teoría tenga valor debe poseer una segunda característica: debe presentar una analogía. Las proposiciones de las hipótesis deben ser análogas a algunas leyes conocidas." Y agrega: "Las analogías no 'ayudan' al establecimiento de teorías; ellas son una parte absolutamente esencial de las teorías, sin las cuales éstas serían totalmente inútiles e indignas de tal nombre".[144] En apoyo de esta afirmación, Campbell construye un pequeño sistema cuasi teórico que implica deductivamente una ley empírica, pero que no es, claramente, una teoría científica aceptable; y esto obedece, en opinión de Campbell, a que su hipótesis carece de la analogía requerida con leyes conocidas. Examinemos brevemente este sistema, al que llamaré *S*.[145]

La hipótesis de *S* está expresada en términos de cuatro conceptos teóricos cuantitativos, *a*, *b*, *c* y *d*, que son funciones de ciertas "variables independientes" *u*, *v* y *w*,... La hipótesis dice que *a* y *b* son funciones constantes, y que *c* es igual a *d*.

El diccionario de *S* consiste en las dos especificaciones siguientes:

142 Campbell (1920, pág. 71).

143 Campbell (1920, pág. 122) dice: "El diccionario relaciona algunas de estas proposiciones, cuya verdad o falsedad se conoce, con ciertas proposiciones que contienen las ideas hipotéticas declarando que si el primer conjunto de proposiciones es verdadero, entonces el segundo conjunto también lo es, y *viceversa*; puede expresarse esta relación diciendo que el primer conjunto *implica* el segundo" (las bastardillas son mías). Evidentemente, éste es un uso no común de la palabra 'implica". En el análisis siguiente, pues, usaré la frase "implica deductivamente" para referirme a la relación lógica no simétrica, en contraposición con la relación simétrica en la que piensa Campbell y que yo indiqué al decir que, según Campbell, el diccionario *traduce* ciertas proposiciones teóricas a proposiciones empíricas.

144 Campbell (1920, pág. 129).

145 Véase Campbell (1920, págs. 123-124).

la afirmación de que $(c^2 + d^2)\ a = R$, donde R es un número racional positivo, implica que la resistencia de un trozo particular de metal puro es R; y la afirmación de que $cd/b = T$ implica que la temperatura del trozo de metal es T.

Ahora bien, la hipótesis de S implica deductivamente que

$$(c^2 + d^2)a \bigg/ \frac{cd}{b} = 2\,ab = \text{constante}$$

Interpretando el cociente de la izquierda mediante el diccionario, obtenemos, según Campbell, la siguiente ley: "La razón de la resistencia de un trozo de metal puro a su temperatura absoluta es constante". (En realidad, esta proposición se sigue sólo para el trozo de metal particular mencionado en el diccionario; pero dejemos esto de lado, ya que no es esencial para la idea que estamos examinando.)

Esta ley, pues, es lógicamente deducible del sistema S y, en este sentido, es explicada por S. Pero Campbell arguye: "Si no se requiriera nada más que esto, nunca nos faltarían teorías que explicaran nuestras leyes; un escolar, en la labor de un día, podría resolver los problemas en los que han trabajado en vano generaciones mediante el más trivial proceso de ensayo y error. Lo erróneo en la teoría..., lo que la hace absurda e indigna aun de ser considerada por un solo momento, es que no presenta ninguna analogía".[146]

Campbell tiene razón, ciertamente, al rechazar la "teoría" S, pero su diagnóstico de sus defectos me parece incorrecto. Lo erróneo en la teoría, creo yo, es que no tiene consecuencias empíricamente testables diferentes de la ley en cuestión (y de todo lo que ella sola implica lógicamente); mientras que una teoría científica útil explica una ley empírica presentándola como un aspecto de regularidades subyacentes más amplias, que tiene también otros diversos aspectos testables, es decir, que también implican diversas leyes empíricas. Tal teoría, pues, suministra una explicación sistemáticamente unificada de muy diferentes leyes empíricas. Además, como observamos en la sección 2, una teoría normalmente implicará refinamientos y modificaciones de leyes empíricas establecidas con anterioridad, en lugar de implicar por deducción las leyes tales como fueron formuladas en su origen.

El diagnóstico de que este defecto, y no la ausencia de analogías, es lo que descalifica a S puede fundamentarse también con la observación de que es posible construir fácilmente sistemas que presenten alguna analogía con leyes conocidas y que, no obstante esto, son inútiles para la ciencia porque tienen el mismo defecto que S. Por ejemplo, sea un sistema S' cuya hipótesis afirma de cuatro magnitudes teóricas, a, b, c y d, que para todo objeto u,

[146] Campbell (1920, págs. 129-130). Campbell admite, sin embargo, que hay un tipo de teoría, ilustrado por la teoría de la conducción del calor de Fourier, en la cual la analogía desempeña un papel menos importante (págs. 140-144). Es claro que para los fines de nuestro examen no necesitamos considerar esas teorías.

$$c(u) = \frac{k_1 a(u)}{b(u)}; \quad d\,(u) = \frac{k_2 b(u)}{a(u)}$$

donde k_1 y k_2 son constantes numéricas; y el diccionario de S' especifica que, para todo trozo u de metal puro, $c(u)$ es su resistencia y $d(u)$ la recíproca de su temperatura absoluta. Entonces, S' también implica deductivamente la ley citada antes y, además, cada una de las dos proposiciones de la hipótesis presenta una analogía con una ley conocida; por ejemplo, con la ley de Ohm . Sin embargo, S' no adquiere más categoría de teoría científica que S, y ello evidentemente por la misma razón.

Así, si bien a mi juicio Campbell no logra demostrar que la analogía desemplea un papel lógico-sistemático esencial en la teorización científica y la eplicación teórica, algunas de sus declaraciones ubican directamente su exigencia de analogías dentro del dominio de los aspectos pragmático-psicológicos de la explicación. Ilustra esto su afirmación de que "una analogía es una función de la mente observadora; cuando decimos que un conjunto de proposiciones es análogo a otro, estamos diciendo algo acerca de su efecto sobre nuestras mentes; produzca o no este efecto sobre las mentes de otros, lo mismo tendrá tal efecto sobre la muestra".[147] Sin duda, concebida de este modo subjetivo, la analogía no puede ser un aspecto indispensable de teorías científicas objetivas.

Si se considera el gran valor heurístico de las analogías estructurales, es natural que un científico que trata de elaborar una nueva teoría se deje guiar por conceptos y leyes que han probado ser fructíferos en dominios ya explorados. Pero si ellos fracasan, tendrá que recurrir a ideas que se aparten cada vez más de las familiares. En la primera teoría del átomo elaborada por Bohr, por ejemplo, la suposición concerniente a la existencia de electrones que giran alrededor del núcleo sin emitir energía viola los principios de la electrodinámica clásica; y en el posterior desarrollo de la teoría cuántica, la analogía de los principios básicos con "leyes conocidas" fue reducida considerablemente, pero la teoría adquirió mayor amplitud, poder explicativo y poder predictivo.

Lo que queda, pues, como principal requisito de la explicación científica es la subsunción inferencial del explanandum bajo principios generales amplios, independientemente de las analogías que puedan presentar con leyes anteriormente establecidas.

Pero hay otro tipo de modelo, llamado a menudo modelo teórico o matemático, que se usa mucho para fines explicativos, por ejemplo en la psicología y la economía. Lo ilustran los abundantes modelos matemáticos del aprendizaje, los modelos teóricos de cambios de actitud y conducta conflictual, y una gran variedad de modelos para fenómenos sociales, políticos y económicos.[148]

[147] Campbell (1920, pág. 144). Más aclaraciones sobre estos problemas se encontrarán en Hesse (1963). El capítulo 2 de este libro tiene la forma de un diálogo entre un "campbelliano" y un "duhemiano", en el que se pasan revista y se evalúan diversos argumentos concernientes a la significación de los modelos y las analogías para la teorización científica.

[148] La bibliografía sobre la cuestión es vasta y aquí sólo podemos dar unas

Hablando en términos amplios, y dejando de lado muchas diferencias de detalle, un modelo teórico de este tipo tiene el carácter de una teoría con un ámbito de aplicación más o menos limitado. Sus suposiciones básicas se refieren a las interdependencias de diferentes características de los fenómenos en cuestión. A menudo, aunque no siempre, esas características están representadas por parámetros o "variables" cuantitativas y pueden ser más o menos directamente observables o medibles, o tener el carácter de conceptos teóricos con una interpretación empírica parcial, por lo menos, efectuada quizá mediante una "definición operacional". Esto es verdad, por ejemplo, acerca de aquellos parámetros que representan probabilidades estadísticas de ciertos tipos de conducta. Las hipótesis básicas del modelo a menudo conciben algunos de los parámetros como funciones matemáticas de otros, pero no siempre tienen este carácter cuantitativo.[149] A partir de las hipótesis básicas, y en conjunción con la interpretación, pueden inferirse consecuencias específicas concernientes a los fenómenos empíricos que el modelo representa: de este modo se hace posible someter a prueba el modelo y darle un uso explicativo y predictivo. Las explicaciones y predicciones resultantes pueden ser deductivo-nomológicas o inductivo-estadísticas, según la forma de la hipótesis incluidas en el modelo. El uso del término "modelo teórico" en lugar de "teoría" quizá pretenda indicar que los sistemas aludidos tienen limitaciones claras, en particular cuando se los compara con teorías físicas avanzadas. Para comenzar, a menudo se sabe que sus suposiciones básicas son idealizaciones o esquematizaciones. Por ejemplo, puede dejar de lado ciertos factores que son de reconocida atingencia a los fenómenos en estudio; esto sucedería, por ejemplo, con un modelo teórico de la conducta económica basado en la suposición de una estricta racionalidad económica por parte de los agentes que intervienen. En segundo término, puede esquematizarse deliberadamente la formulación de las interrelaciones entre diferentes factores, quizá con el fin de facilitar matemáticamente la aplicación del modelo a casos particulares. Además, la clase de fenómeno que el modelo representa puede ser muy limitada; por ejemplo, un modelo teórico de la adopción de decisiones en situaciones de riesgo podría restringirse a las decisiones que se

---

pocas referencias específicas. Un examen general particularmente claro de los modelos teóricos en psicología, además de un modelo específico de conducta conflictiva, se expone en Miller (1951). Sobre modelos del aprendizaje véase, por ejemplo, Bush y Mosteller (1955); en la introducción de este libro se aclara la metodología del procedimiento utilizado por el autor. El volumen colectivo de Lazarsfeld (1954) contiene la exposición de modelos matemáticos para diversos aspectos de la conducta social, así como ensayos dedicados al análisis de modelos particulares o a problemas generales concernientes a la metodología de la construcción de modelos. Una excelente exposición general del papel de los modelos matemáticos en las ciencias sociales se ofrece en Arrow (1951), y los simposios de la Sociedad de Biología Experimental (1960) y de la Unión Internacional de Historia y Filosofía de las Ciencias (1961) contienen algunos artículos interesantes sobre el papel de los modelos en las ciencias empíricas. El ensayo de Brodbeck (1959) presenta observaciones esclarecedoras sobre el carácter y la función de los modelos teóricos.

[149] Esto es verdad, por ejemplo, del modelo teórico de Miller de la conducta conflictual, formulado en términos de hipótesis comparativas tales como "la tendencia a lograr un objetivo es más fuerte cuanto más cerca de él está el sujeto". Miller (1951, pág. 90).

toman en condiciones más bien artificiales experimentalmente controladas, y que se limitan a un número pequeño de opciones bastante triviales.

Pero tales peculiaridades pueden hallarse también en el campo de la teorización física, y ellas no quitan a los sistemas aludidos el carácter de teorías potencialmente explicativas. Sin embargo, un ámbito limitado y una validez sólo aproximada dentro de este ámbito pueden disminuir mucho el valor explicativo y predictivo real de un modelo teórico.

7. *Explicación genética y leyes inclusivas*

Los modelos de leyes inclusivas a menudo han sido criticados sobre la base de que si bien pueden representar correctamente la estructura y el alcance de algunas de las explicaciones formuladas por las ciencias empíricas, no hacen justicia a muchas otras. En esta sección y en las que siguen me propongo examinar algunos modos y aspectos importantes de la explicación científica que se han citado en apoyo de esta afirmación, y trataré de señalar cómo el concepto de ley inclusiva puede aclarar su lógica y su fuerza.

Un procedimiento explicativo muy utilizado en la historia, aunque no solamente en ésta, es el de la explicación genética. Ella presenta el fenómeno en estudio como la etapa final de un desarrollo, y explica dicho fenómeno describiendo las etapas sucesivas de este desarrollo.

Consideremos, por ejemplo, la costumbre de vender indulgencias en la forma adoptada cuando Lutero era joven. El historiador eclesiástico H. Boehmer nos dice que hasta comienzos del siglo XX "la indulgencia todavía era, en realidad, una magnitud desconocida, ante la cual el erudito se preguntaba con un suspiro: '¿de dónde provino?' " Adolf Gottlob sugirió una respuesta; abordó el problema preguntándose qué condujo a los papas y obispos a otorgar indulgencias. Como resultado de esto "...el origen y el desarrollo de esa magnitud desconocida apareció claramente a la luz y terminaron las dudas acerca de su significado original. Se reveló como una herencia de la época de la gran lucha entre el Cristianismo y el Islam, a la par que un producto muy característico de la cristiandad germánica".[150]

Según esta tesis[151] los orígenes de la indulgencia se remontan al siglo IX, cuando la lucha contra el Islam era una preocupación importante de los papas. Las enseñanzas de su religión aseguraban al guerrero mahometano que si moría en la batalla su alma iba inmediatamente al cielo, pero el cristiano debía temer su perdición si no había hecho la penitencia regular por sus pecados. Para aliviar estas dudas, Juan VII, en 877, prometió la absolución de pecados a los cruzados que murieran en combate. "Una vez jerarquizada de este modo la cruzada, fue fácil pasar a considerar la par-

[150] Boehmer (1930, pág. 91). El estudio de Gottlob, *Kreuzablass und Almosenablass*, fue publicado en 1906; véase las referencias a la obra de Gottlob y de otros investigadores en Schwiebert (1950), notas al capítulo 10.

[151] Sigo aquí las exposiciones de Boehmer (1930, capítulo 3) y Schwiebert (1950, capítulo 10).

ticipación en una cruzada como equivalente a la ejecución de una expiación... y a prometer la remisión de... las penitencias a cambio de realizar expediciones contra los enemigos de la Iglesia".[152] Así se introdujo la indulgencia de la Cruz, que otorgaba la remisión completa de los castigos penitenciales a quienes participasen en una guerra religiosa. "Si se recuerda cuántos inconvenientes, cuántas desventajas eclesiásticas y civiles implicaban las penas eclesiásticas, fácil es comprender que los penitentes se congregaran para obtener esta indulgencia".[153] Otro fuerte incentivo fue la creencia de que quien obtenía una indulgencia, no sólo se aseguraba la exención de las penitencias eclesiásticas, sino también de los correspondientes sufrimientos en el Purgatorio después de la muerte. Los beneficios de estas indulgencias se extendieron luego a aquellos que, físicamente incapacitados para participar en una guerra religiosa, contribuían a reunir los fondos necesarios para enviar a un soldado a una cruzada. En 1199, el Papa Inocencio III admitió el pago de dinero como una condición adecuada para gozar de los beneficios de una indulgencia referente a una cruzada.

Cuando las cruzadas declinaron, se exploraron nuevos medios para obtener fondos mediante las indulgencias. Así, se instituyó una "indulgencia de jubileo" que se celebraba cada cien años, en beneficio de los peregrinos que acudieran a Roma en esta ocasión. La primera de estas indulgencias, en 1300, produjo grandes sumas de dinero, y el intervalo entre sucesivas indulgencias de jubileo fue reduciéndose a 50, 33 y aun 25 años. A partir de 1393, pudo disponerse de la indulgencia de jubileo no sólo en Roma, sino en toda Europa, por medio de agentes especiales que fueron autorizados a absolver a pecadores penitentes mediante un pago adecuado. La evolución de esta costumbre dio un paso más: en 1477, una declaración dogmática de Sixto IV atribuyó a la indulgencia el poder de liberar hasta a los muertos del Purgatorio.

Es innegable que una explicación genética de este tipo puede aclarar nuestra comprensión de un fenómeno histórico. Pero yo creo que su función explicativa tiene un carácter básicamente nomológico, pues las etapas sucesivas señaladas para su examen interesan por su función, más que por el hecho de formar una sucesión temporal y de preceder todas ellas a la etapa final, que es la que se debe explicar: la mera enumeración en un anuario de los "sucesos importantes del año" en el orden en que se produjeron, evidentemente no es una explicación genética del suceso final ni de ninguna otra cosa. En una explicación genética, debe mostrarse que cada etapa "conduce" a la siguiente, y de este modo, que se vincula con ésta en virtud de algunos principios generales que hacen razonablemente probable, al menos, la aparición de ella, dada la primera. Pero en este sentido, aun las etapas sucesivas de un fenómeno físico, tal como la caída libre de una piedra, pueden considerarse como si constituyeran una sucesión genética cuyas diferentes etapas —caracterizadas, digamos, por la posición y la velocidad de la piedra de diferentes momentos— se hallan interconectadas por leyes estrictamente universales; y las etapas sucesivas

152 Boehmer (1930, pág. 92).
153 Boehmer (1930, pág. 93).

de una bola de acero que rebota en zig zag hasta un tablero de Galton[154] pueden considerarse como si formaran una sucesión genética con vínculos probabilísticos.

Las explicaciones genéticas dadas por los historiadores no son, por supuesto, del tipo puramente nomológico sugerido por estos ejemplos tomados de la física. Más bien combinan un cierto grado de interconexión nomológica con grados más o menos grandes de descripción directa. Por ejemplo, consideraremos una etapa intermedia mencionada en una explicación genética. Algunos de sus aspectos se presentarán como una evolución desde las etapas anteriores (en virtud de leyes de conexión, las cuales a menudo sólo serán mencionadas); otros aspectos, que no quedan explicados por la información que se posee acerca del desarrollo anterior, se agregarán descriptivamente porque son importantes para comprender etapas posteriores de la sucesión genética. Así, hablando de manera esquemática, una explicación genética comenzará con una descripción pura de una etapa inicial; de aquí pasará a una explicación de la segunda etapa, parte de la cual se halla vinculada nomológicamente con las características de la etapa inicial y es explicada por éstas, a la par que se agrega el resultado total de manera simple y descriptiva, por su importancia para la explicación de algunas partes de la tercera etapa, y así sucesivamente.[155]

El siguiente diagrama representa de un modo esquemático la manera cómo la explicación nomológica se combina con la descripción directa en una explicación genética del tipo considerado:

$$S_1 \left.\begin{matrix} \nearrow S'_2 \\ \\ +D_2 \end{matrix}\right\} S_2 \left.\begin{matrix} \nearrow S'_3 \\ \\ +D_3 \end{matrix}\right\} S_3 \ldots \left.\begin{matrix} \nearrow S'_{n-1} \\ \\ +D_{n-1} \end{matrix}\right\} S_{n-1} \rightarrow S_n$$

Cada flecha indica una presunta conexión nómica entre dos etapas sucesivas; propone uniformidades que, por lo general, no se enuncian de manera completa y explícita, y que pueden ser del tipo estrictamente universal o —con mayor probabilidad— de tipo estadístico. O_1, O_2, ..., O_n son conjuntos de oraciones que expresan toda la información que la explicación genética da acerca de la primera, la segunda, ..., la enésima etapa. Para cada una de estas etapas, excepto la primera y la última, la información así suministrada se divide en dos partes: una de ellas —representada por S'_2, S'_3, ..., S'_{n-1}— describe aquellos hechos de la etapa considerada que se explican con referencia a la etapa anterior; la otra —representada

154 Una descripción del mecanismo y un análisis probabilístico de sus operaciones se hallarán, por ejemplo, en Mises (1939, págs. 237-240).

155 Esta concepción de la estructura de la explicación genética en la historia está básicamente de acuerdo con la expuesta por Nagel (1961, págs. 564-568) en el contexto de un análisis muy sustancial y amplio de los problemas de la lógica de la investigación histórica. También Frankel (1959, pág. 412) y Goldstein (1958, págs. 475-479) destacan la presuposición de generalizaciones vinculadoras. Sobre el papel de la "narración coherente" en contraposición a la explicación por leyes inclusivas en la historia natural véase también Goudge (1958).

por D_2, D_3, ..., D_{n-1}— consiste en una información acerca de otros hechos que se presenta sin explicación, por su importancia explicativa para la etapa siguiente. Apenas será necesario señalar que esta caracterización de la explicación genética es muy esquemática; ella pretende mostrar las afinidades que tiene este procedimiento con la explicación nomológica, por una parte, y con la descripción, por la otra. En la práctica, a menudo será difícil separar estos dos componentes: en lugar de presentar con nitidez un conjunto de etapas interconectadas pero distintas en una sucesión temporal, es probable que una explicación genética suministre descripciones de una gran variedad de hechos y sucesos que se extienden en un cierto ámbito temporal y no son fácilmente agrupables en cúmulos que constituyan etapas sucesivas, y sugiera conexiones entre tales hechos y sucesos.

En nuestro ejemplo, se indica la suposición de ciertas leyes o principios de carácter legal que vinculan los hechos mediante referencias o factores motivacionales. Así, las explicaciones dadas para el deseo de los papas de organizar una fuerza de combate o reunir fondos aun mayores presuponen claramente tesis psicológicas acerca de la manera en que tenderá a actuar un individuo inteligente, a la luz de sus creencias reales, cuando trata de alcanzar determinado objetivo. Las uniformidades psicológicas se hallan también implícitas en la referencia al temor al Purgatorio para explicar la ansiedad con la cual se compraban las indulgencias. Igualmente, cuando un historiador observa que el enorme éxito financiero de la primera indulgencia de jubileo "no hizo más que aguzar el insaciable apetito de los papas" y que "el período se redujo de 100 a 50, a 33 y a 25 años",[156] la explicación sugerida reposa en una suposición psicológica afín a la idea de refuerzo por recompensas. Pero, por supuesto, aun cuando se adujera explícitamente alguna formulación de esta idea, la explicación resultante brindaría a lo sumo una explicación parcial; por ejemplo, no podría explicar por qué los intervalos sucesivos tuvieron las particulares longitudes mencionadas.

Estos factores, a los que en nuestro ejemplo simplemente se describe o se los presupone de modo tácito como "hechos en bruto", para usar la expresión de Nagel,[157] incluyen, por ejemplo, las doctrinas, la organización y el poder de la Iglesia, el surgimiento de las cruzadas y la eventual decadencia de este movimiento, y muchos factores adicionales que no se mencionan de modo explícito, pero que deben comprenderse como condiciones de fondo para que la explicación genética cumpla su finalidad explicativa.

Consideremos brevemente otro ejemplo de explicación genética tomado de Toynbee. En 1839 la principal maternidad de la ciudad de Alejandría fue ubicada en los terrenos del arsenal naval. "Esto suena extraño —observa Toynbee— pero veremos que era inevitable tan pronto como sigamos el desarrollo de la serie de sucesos que condujo a este resultado en primera instancia sorprendente".[158] La explicación genética de Toynbee es, brevemente, la siguiente. Por el año 1839 hacía ya más de 30 que Mohamed 'Ali Pasha, gobernador otomano de Egipto, trataba de equipar-

[156] Schwiebert (1950, pág. 304).
[157] Nagel (1961, pág. 566).
[158] Toynbee (1953, pág. 75).

se con armamentos efectivos, en particular con una flota de barcos al estilo occidental. Comprendió que su establecimiento naval no tendría autonomía a menos que estuviera en condiciones de construir sus barcos de guerra en Egipto con trabajadores nativos, y que un grupo eficiente de técnicos navales egipcios sólo podía ser preparado por especialistas occidentales, quienes tendrían que ser contratados para este propósito. El gobernador, pues, publicó anuncios para atraer expertos occidentales, ofreciéndoles salarios muy tentadores. Pero los especialistas que se presentaron no querían ir sin sus familias, y deseaban asegurarse una atención médica que estuviera en el nivel de las normas occidentales para que atendieran a los expertos navales y sus familias. Pero los médicos descubrieron que tenían tiempo para realizar una labor adicional; "eran médicos activos, dinámicos y de espíritu social, por lo que resolvieron hacer algo también por la población egipcia local... la primera exigencia obvia era la maternidad. Así surgió una maternidad dentro del recinto del arsenal naval, por una serie de sucesos que, como ahora reconoceréis, era inevitable".[159]

Toynbee trata así de explicar el hecho en cuestión, al principio extraño, mostrando cómo se produjo "inevitablemente" como etapa final de una sucesión de acontecimientos interconectados; y se refiere a este caso como un ejemplo del "proceso de una cosa que conduce a otra" [160] en las relaciones interculturales. Pero, ¿dónde reside la inevitabilidad con la cual una cosa conduce a la siguiente? En varios puntos de la explicación de Toynbee se sugiere la presunta conexión mediante la referencia explicativa a las razones motivacionales de los agentes; pero éstas sólo brindan fundamentos explicativos para las acciones resultantes en la suposición de que las personas impulsadas por tales y cuales razones *generalmente* actuarán, o *tenderán* a actuar, de ciertas maneras características. Por ende, la idea de una cosa que conduce inevitablemente a otra presupone aquí una conexión mediante principios con carácter de leyes que son válidas para ciertos tipos de acciones humanas. El carácter de tales principios y la lógica de las explicaciones basadas en ellos serán examinados con más detalle en las secciones 9 y 10 de este capítulo.

Ahora consideraremos brevemente algunos puntos controvertidos acerca de las explicaciones genéticas en la historia y sobre los cuales pueden arrojar alguna claridad las consideraciones anteriores.

Dray ha sostenido que la explicación genética en historia tiene peculiaridades lógicas que pueden ponerse de manifiesto mediante una comparación con lo que llama "el modelo de la serie continua".[161] Ilustra este modelo mediante una explicación de la detención de un motor de automóvil que la atribuye a una pérdida de aceite: como resultado de la pérdida, el aceite se agota, privando de lubricación a los cilindros y los pistones, lo cual provoca el calentamiento por fricción y la dilatación de los pistones y las paredes de los cilindros, de modo que los metales se traban y el motor se detiene. Dray pone mucho énfasis en la afirmación de que al revelar el mecanismo de la falla, esta explicación gradual suministra una compren-

159 *Ibíd.*, pág. 77.
160 *Ibíd.*, pág. 75.
161 Dray (1957, págs. 66 y sigs.).

sión que no brindaría una ley inclusiva que vinculara la falla directamente con la pérdida: "*Por supuesto,* el motor se detiene, y digo esto porque ahora puedo considerar una *serie continua de sucesos* entre la pérdida y la detención del motor que son en sí mismos claramente comprensibles, mientras que la sucesión original 'pérdida-a-detención' no lo era".[162]

Si la entiendo con corrección, la defensa de Dray de esta afirmación se basa en considerable medida en innegables diferencias pragmáticas entre las dos explicaciones: la explicación gradual suministra una comprensión que no se logra cuando se vinculan de manera inmediata las etapas final e inicial. Pero esta diferencia pragmática está asociada, creo, a otra no pragmática que justifica la aserción de que las dos explicaciones difieren en poder explicativo. Para comprender esto, otorguemos carácter nomológico —para los fines de la argumentación— al enunciado *L*, según el cual siempre que la caja del aceite de un automóvil adecuadamente construido tiene una pérdida, su motor falla. Esta ley podría ser invocada, entonces, para dar una explicación de bajo nivel de ciertos casos particulares de fallas de los motores. La explicación gradual, en cambio, presenta el proceso a través de una sucesión de etapas, y a cada una de éstas como gobernadas por ciertas "subleyes", como Dray las llama, por ejemplo las que vinculan la fricción entre los pistones y las paredes del cilindro con el calentamiento y la dilatación de los metales. Pero un conjunto adecuado de tales leyes no sólo nos permitirá explicar casos particulares de fallas de motores, sino también explicar por qué la ley *L* se cumple, esto es, por qué una pérdida de aceite en un automóvil bien construido *generalmente* provoca una falla del motor.

En el caso de la explicación genética en historia, hay una razón adicional para pensar que una explicación por etapas es esencial para lograr la comprensión: aquí no tenemos leyes generales que, en analogía con la ley *L* del ejemplo anterior, vinculen inmediatamente la etapa final del proceso con la inicial. Como indica nuestra caracterización esquemática, los datos particulares acerca de la etapa inicial no bastan por sí mismos para explicar todos los aspectos específicos de la etapa final. Para explicarlos, necesitamos datos adicionales, y éstos los suministra poco a poco la información acerca de "hechos en bruto" adicionales que figuran en las descripciones de las etapas intermedias.

Nuestra concepción de la explicación genética también toma en cuenta la queja de que las leyes que podríamos aducir realmente en el contexto de la explicación histórica, incluso leyes psicológicas y otras de la experiencia común, resultan triviales e inadecuadas cuando tratamos de dar cuenta de las ricas y características peculiaridades que presuntamente dan un carácter único a los sucesos históricos y que son, por lo tanto, de especial interés para el historiador. Si se considera, por ejemplo, la sutileza y complejidad de algunas de las explicaciones psicológicas que se han propuesto con referencia a las acciones de personajes históricos, esta acusación quizá sea un tanto exagerada; pero es innegable que tiene un valor considerable. Y el modelo que acabamos de esbozar tiene en cuenta esta dificultad pre-

162 *Ibid., pág.* 68; bastardillas del autor. Observaciones con un espíritu similar y otros ejemplos se hallarán en Danto (1956, págs. 23-25).

viendo la introducción en una explicación genética de una masa más o menos grande de detalles, los que simplemente describe, sin explicarlos, haciendo referencia a otros hechos particulares y uniformidades vinculantes.

8. *Explicación por conceptos*

Otro modo de explicación que presumiblemente presenta dificultades para la concepción basada en leyes inclusivas ha sido señalado por Dray, quien considera su papel en la investigación histórica. Dray la llama "explicar lo que..." o "explicación por conceptos", sobre la base de que un pedido de explicación de este tipo adquiere típicamente la forma "¿qué fue lo que sucedió en este caso?", y de que el historiador "la aborda ofreciendo una explicación de la forma 'fue tal y tal cosa' ".[163] Dray ilustra la idea tomada de la obra de Ramsey Muir *Short History of the British Commonwealth* [Breve historia del Commonwealth británico]. Este describe ciertos cambios que se produjeron a fines del siglo XVIII en Inglaterra —tales como el cercamiento de tierras agrícolas, los comienzos de la producción industrial y el mejoramiento de las comunicaciones— y luego continúa: "No era solamente un cambio económico el que se estaba iniciando de este modo; era una revolución social". Dray arguye que si bien el historiador no pretende decirnos aquí por qué o cómo se produjeron los sucesos investigados, su "afirmación de que 'era una revolución social' es, sin embargo, una explicación. Explica lo que sucedió *como* una revolución social".[164] Dray caracteriza este tipo de explicación como una "explicación por medio de un concepto general, no por una ley general; pues se da la explicación apelando a una *clasificación* satisfactoria de lo que parece requerir explicación".[165] Dray agrega que si alguna generalización es esencial para este tipo de explicación, ella no adopta la forma de una ley general. En efecto: "Lo que se quiere explicar es un conjunto de sucesos o condiciones, x, y y z; y la generalización pertinente sería de la forma: 'x, y y z equivalente a Q'. Tal generalización explicativa es sumaria; ella nos permite referirnos a x, y y z colectivamente como a 'tal y tal cosa'. Y los historiadores hallan satisfactorio intelectualmente poder representar los sucesos y condiciones que estudian como si estuvieran relacionados de esta manera".[166]

Pero, sin duda, no toda explicación semejante puede ser considerada como explicativa: los sucesos particulares a los que alude Muir, por ejemplo, podrían clasificarse también, de manera fiel pero poco esclarecedora, como cambios en los que intervienen más de 1000 personas y abarcan una superficie de más de 259 kilómetros cuadrados. Si tiene significación explicativa caracterizar colectivamente a x, y y z como un Q, es porque la

[163] Dray (1959, pág. 403); bastardillas del autor.
[164] *Ibíd.*; bastardillas del autor.
[165] Dray (1959, pág. 404); bastardillas del autor.
[166] *Ibíd.*, pág. 406.

caracterización implica que los casos particulares se adecuan o ajustan a algún esquema general que es característico de *Q*.

Ilustraré lo anterior, primero, mediante algunos ejemplos que revelan, al mismo tiempo, que dicho procedimiento también se usa fuera del dominio de la historiografía.

La explicación de Torricelli de por qué una simple bomba aspirante puede elevar agua a no más de 10 metros se basa, según se ha dicho, en el "esquema conceptual" que postula la existencia de un "mar de aire" que rodea a la Tierra.[167] Pero es claro que este esquema sólo tiene fuerza explicativa porque supone una analogía nómica entre el mar de aire y un mar de agua, a saber, que "habrá una presión del aire sobre todos los objetos sumergidos en este mar de aire, exactamente como hay una presión del agua por debajo de la superficie del océano",[168] y que la presión está determinada por el peso de la columna de aire que está por encima del objeto mencionado. Así fue, en efecto, como razonó Torricelli. Luego, la explicación por medio de este esquema conceptual efectúa una subsunción del fenómeno del explanandum bajo hipótesis generales.

Ahora, como ejemplo que presenta una clara semejanza con el citado por Dray, consideremos el siguiente enunciado: "El catarro nasal y los ojos inflamados de Otto, así como las manchas rojas rodeadas de superficies blancas que acaban de aparecer en las mucosas interiores de sus mejillas no son sucesos aislados; son, todos ellos, síntomas que señalan el comienzo de un caso típico de sarampión". Esta clasificación diagnóstica explica los síntomas particulares citados señalando que, en conjunto, se conforman al cuadro clínico del sarampión; es decir que se presentan ciertos tipos característicos y aparecen en un orden temporal característico, que serán seguidos por otros síntomas específicos y que la enfermedad tenderá a seguir un curso característico. Interpretar un conjunto de males como manifestación del sarampión equivale, sin duda, a sostener que ellos entran en determinado esquema de regularidades (que serán de tipo estadístico, más que de forma estrictamente universal) ; y tal explicación concuerda con la concepción basada en leyes inclusivas de la explicación.

O bien consideremos la "clasificación" de una sucesión particular de rayos y truenos como un caso de potente descarga eléctrica que genera una violenta perturbación del aire. Esto tiene, en efecto, contenido explicativo, pero porque señala que el particular conjunto de sucesos presenta las características que manifiestan generalmente las descargas potentes y las perturbaciones que provocan en el aire; o con mayor precisión, ellas se conforman a las leyes características del tipo de fenómeno en el que se clasifica el caso particular considerado.

En la cita que hace Dray de Muir, la declaración "era una revolución social" transmite, análogamente, la sugerencia de que se brinda un diagnóstico explicativo, sugerencia que se refuerza con el siguiente pasaje amplificativo, que sigue de inmediato al pasaje citado por Dray: "El viejo orden asentado y estable que hemos descripto para la Gran Bretaña de mediados del siglo XVIII se estaba transformando de manera total... Pero

[167] Conant (1951, pág. 69).
[168] *Ibíd.*

todavía no se comprendía la plena significación de este cambio. Tranquilamente asentadas en el poder, las viejas clases gobernantes mostraban una completa ceguera ante las fuerzas que operaban por debajo de sus pies socavando los cimientos mismos de su poder y haciendo inevitable que tarde o temprano el sistema político debiera reajustarse de acuerdo con los cambios del orden social".[169] Tenemos aquí la sugerencia de un diagnóstico, o interpretación, según el cual los cambios particulares en la agricultura, la producción industrial y las comunicaciones que Muir había descripto antes eran las tempranas manifestaciones de un proceso más amplio, cuyas diferentes fases no estaban asociadas de manera casual, sino con cierto carácter de inevitabilidad. Así, nuevamente, aunque sólo de una manera vaga y muy esquemática, se asigna a los casos particulares un lugar en una vasta trama de conexiones. Sea cual fuere el valor explicativo que pueda tener la declaración de Muir —que, para mí, es escaso—, él reside, sin duda, en sugerir un diagnóstico del tipo ilustrado con más claridad por los dos casos anteriores, que se conforman, en líneas generales, a la concepción basada en leyes inclusivas.

Otros ejemplos de lo que Dray llama explicación-por-conceptos los suministran las diversas interpretaciones de la Guerra de Secesión norteamericana como resultado de una conspiración de algunos grupos norteños, o sureños, de "hombres malvados", como una querella entre dos regiones rivales, como una disputa sobre tipos de gobierno, como una consecuencia del "irrefrenable conflicto" entre la libertad y la esclavitud, como un antagonismo básicamente económico, etc.[170] Cada una de estas explicaciones de la Guerra de Secesión "como un tal y tal" atribuye significación causal especial o predominante a factores de algún tipo especial y, por consiguiente, presupone adecuadas conexiones nómicas en apoyo de tales afirmaciones.[171]

Dray reconoce explícitamente que "la explicación-por-conceptos a veces puede subsumir *de hecho* el explicandum bajo una ley",[172] pero sostiene que esto no es lo general. De manera específica, está en desacuerdo con mi anterior afirmación según la cual "lo que a veces se llama, equivocadamente, una explicación por medio de cierto *concepto* es en realidad, en las ciencias empíricas, una explicación en función de *hipótesis universales* que contienen dicho concepto".[173] En contra de esta idea, Dray arguye del siguiente modo: "Se presume que la ley que subyace en el fondo cuando se explica 'algo como una revolución' es una ley que contiene el

169 Muir (1922, pág. 123).

170 Sobre estas diferentes interpretaciones véase, por ejemplo, Beale (1946).

171 El problema de evaluar factores causales según su importancia relativa en una explicación histórica está tratado por Nagel (1961, págs. 582-588) con claridad.

172 Dray (1959, pág. 405); bastardillas del autor.

173 Hempel (1942, nota 3); bastardillas del original. Homans ha hecho recientemente la misma observación con referencia a la sociología. Sostiene que mucho de la teoría sociológica moderna no explica nada, en parte porque "consiste en gran medida en sistemas de categorías, o casilleros, en los que el teórico ubica diferentes aspectos de la conducta social ... pero esto no basta para darle poder explicativo ... La ciencia también necesita un conjunto de proposiciones generales acerca de las relaciones entre las categorías, pues sin tales proposiciones la explicación es imposible". Homans (1961, pág. 10).

concepto en su apódosis... Pero explicar, por ejemplo, lo que sucedió en Francia en 1789 'como una revolución' no equivale, sin duda, a una ley de la forma 'siempre que C_1, C_2, ..., C_n, entonces una revolución' ".[174] Pero mi observación anterior no limita una explicación-por-conceptos a una hipótesis general, ni limita las hipótesis explicativas al tipo considerado por Dray. Se aplica también, por ejemplo, a la explicación de ciertos malestares "como síntomas del sarampión", la cual se basa en hipótesis generales que afirman que si una persona sufre de sarampión, entonces presentará síntomas de tal y tal tipo; en este caso, se alude al concepto explicativo en la prótasis, no en la apódosis.

O bien consideremos lo que podría llamarse vagamente la "explicación del brillo de un meteorito que cae como un caso de calor intenso generado por la fricción". Aquí están implicadas varias leyes, entre ellas dos según las cuales un cuerpo que se desplaza por el aire sufre una fricción y esta fricción genera calor. De este modo, podría decirse que los conceptos explicativos figuran en parte en la prótasis y en parte en la apódosis de las correspondientes leyes generales.

El ejemplo de Dray está formulado tan esquemáticamente que es difícil apreciar la explicación que se presume lograda. Un enunciado que caracterice lo sucedido en Francia en 1789 como una revolución parecería suministrar una descripción muy vaga de esos sucesos, más que una explicación de ellos. Podría atribuírsele cierto contenido explicativo si se entendiera el concepto de revolución en un sentido técnico restringido, que implique quizás una serie de etapas características del proceso, ciertos cambios característicos en la estructura del poder político o alguna otra cosa semejante. Entonces podría demostrarse que algunos de los sucesos particulares de 1789 se adecuan a las pautas implicadas por el concepto dado de revolución y, de tal modo, considerarlos como parcialmente explicados por éste. Pero en tal caso, la explicación se lograría, es evidente, mediante una referencia a las uniformidades implicadas.

En resumen, pues, un uso explicativo de conceptos debe basarse siempre en hipótesis generales correspondientes a ellos.

9. *La explicación disposicional*

Otro tipo de explicación del que se ha sostenido que no entra en el análisis basado en leyes inclusivas aduce, de una manera característica, ciertas propiedades disposicionales de los objetos o agentes cuya "conducta" debe ser explicada. A este procedimiento lo llamaré explicación disposicional.

El método corriente para explicar decisiones y acciones humanas en función de propósitos, creencias, rasgos de carácter, etc., es básicamente de este tipo, pues atribuir a un agente tales factores motivantes equivale a asignarle ciertas características disposicionales más o menos complejas.

[174] Dray (1959, pág. 404).

Ryle[175] ha sostenido este concepto, cuyas ideas han ejercido una gran influencia sobre la discusión de este punto. Las explicaciones por razones motivantes serán examinadas con cierto detalle en la sección 10. En esta sección, consideraremos la estructura lógica de algunas explicaciones disposicionales de la física, y la compararemos con la de las explicaciones por leyes inclusivas.

Consideremos primero el ejemplo que analiza Ryle. Cuando el vidrio de un ventana se hace pedazos después de haber sido golpeado por una piedra, la ruptura del vidrio puede explicarse causalmente, según Ryle, señalando que lo golpeó una piedra; pero, a menudo, buscamos una explicación en un sentido diferente: "Preguntamos por qué el vidrio se rompió cuando lo golpeó la piedra y se nos responde que fue debido a que el vidrio es frágil".[176] Aquí se obtiene la explicación, no especificando un suceso independiente "que esté con respecto a la ruptura del vidrio como causa con respecto al efecto",[177] sino atribuyendo al vidrio una cierta propiedad disposicional, la fragilidad. Atribuir esta propiedad a un vidrio de ventana particular equivale, al menos por implicación, a afirmar una hipótesis *general* que diría, aproximadamente, que si se golpea con fuerza en algún momento el vidrio con un cuerpo o si algún agente lo dobla, quedará reducido a pedazos. Pero aunque un enunciado disposicional sea, pues, de carácter general, sin embargo, también menciona un ente particular, como el vidrio de la ventana. A este respecto, los enunciados disposicionales difieren de las leyes generales, que —según Ryle— no mencionan individuos en absoluto. Para indicar sus semejanza con las leyes generales y también su diferencia de ellas, Ryle llama a los enunciados disposicionales "enunciados de tipo legal" (*law-like*).[178]

Debe observarse, sin embargo, que ninguno de los tipos de explicación que distingue Ryle es suficiente por sí mismo para dar cuenta del hecho dado. El informe de que el vidrio fue golpeado por una piedra

[175] Véase en particular Ryle (1949).

[176] Ryle (1949, pág. 88).

[177] *Ibíd.*

[178] Para mayores detalles, véase Ryle (1949, págs. 43-44, 89, 120-125). Hablando en términos estrictos, la distinción propuesta entre oraciones legales y leyes generales no puede elucidarse satisfactoriamente sobre la base de que las oraciones en cuestión "'mencionen o no cosas o personas particulares", como dice Ryle (*loc. cit.*, pág. 123), pues puede evitarse la mención explícita de un individuo mediante una reformulación. Por ejemplo, la oración general "todos los lugares de la superficie de la Tierra que están dentro de los 160 kilómetros del Polo Norte son fríos" sería considerada legal porque menciona al Polo Norte. Sin embargo, se la puede reformular del siguiente modo: "todos los lugares polares son fríos", donde se usa "polar" como sinónimo de "estar en la superficie de la Tierra dentro de los 160 kilómetros del Polo Norte"; y según el criterio considerado, la reformulación debe considerarse como una ley general porque no menciona (es decir, no contiene ninguna designación) a una persona, lugar o cosa particular. Para un examen más completo del problema, véase Hempel y Oppenheim (1948, sección 6) y Goodman (1955, en particular los capítulos 1 y 3). Obsérvese de paso que Goodman usa el término "legal" en un sentido muy diferente que Ryle, o sea, para referirse a oraciones que tienen todas las características de una ley, pero que pueden ser falsas (*loc. cit.*, pág. 27). Para evitar una extensa digresión, renunciaremos aquí a tratar de ofrecer una elucidación más adecuada de la importante distinción establecida por Ryle y consideraremos la idea como suficientemente clara para nuestros fines.

explica su ruptura sólo en conjunción con la información adicional de que el vidrio era frágil: ser golpeado por la piedra es una causa, y no un antecedente accidental, de la ruptura del vidrio en virtud de la hipótesis general implicada por aquella atribución disposicional. De manera análoga, el enunciado disposicional puede explicar la ruptura del vidrio sólo si lo toma juntamente con el informe de que el vidrio es golpeado con fuerza: y en realidad, como vimos, el mismo Ryle describe el enunciado disposicional como explicando "por qué el vidrio se rompió cuando fue golpeado por la piedra" y no simplemente por qué el vidrio se rompió aquí. Así, cualquiera de las dos explicaciones distinguidas es incompleta y exige que se la complemente con la otra. Conjuntamente, suministran una explicación adecuada que podría formularse esquemáticamente del siguiente modo:

(9.1)

(C_1) El vidrio fue golpeado con fuerza por una piedra en el tiempo t_1.

(L_1) Para cualquier tiempo t sucede que si se golpea con fuerza el vidrio en t, éste se rompe en t.

(E_1) El vidrio se rompió en t_1.

Esta explicación es deductivo-nomológica, excepto para aducir un enunciado de tipo legal, en lugar de una ley completamente general. En este aspecto, el razonamiento se halla en buena compañía; las leyes de Galileo y de Kepler, por ejemplo, se usan, sin duda, con propósitos explicativos; sin embargo, la primera, cuando se la enuncia de manera completa, especifica que su fórmula se aplica a la caída libre cercana a la superficie de la Tierra, es decir, menciona un objeto particular; y las leyes de Kepler, tales como fueron concebidas originalmente, se referían a los movimientos de los planetas de un objeto particular, el Sol. Claro que estas leyes han sido subsumidas desde entonces en las leyes newtonianas del movimiento y la gravedad, que son de forma completamente general. Un caso similar es posible en el ejemplo de la ventana rota, donde el enunciado "el vidrio era frágil" puede reemplazarse en el razonamiento explicativo por una hipótesis absolutamente general, "todo vidrio es frágil (en condiciones comunes)", y el enunciado singular "la hoja de la ventana era de vidrio (y estaba en condiciones comunes)".

Sin embargo, las teorías corrientes no nos permiten realizar esta especie de subsunción en leyes o principios teóricos estrictamente generales para todos los enunciados de tipo legal, en particular para todos los enunciados que atribuyen a individuos disposiciones psicológicas. Pero aun en estos casos siempre puede darse otro paso: en lugar de dar al enunciado disposicional explicativo la forma de una generalización que mencione un individuo particular a la manera de L_1 en (9.1), podemos expresarlo mediante dos enunciados separados, uno singular, que afirme que el individuo dado tiene la propiedad disposicional en cuestión, digamos D, y otro completamente general que caracterice la disposición D. En el caso de (9.1), esto sería equivalente a reemplazar la oración L_1 por las dos siguientes:

(C_2) la hoja de la ventana era frágil en el tiempo t_1.

(L_2) todo objeto frágil, si es golpeado con fuerza en cualquier momento, se rompe en este momento.

Podría objetarse que el único enunciado general que aparece en la modificación resultante de (9.1), a saber L_2, no tiene el carácter de una *ley empírica* acerca de objetos frágiles, sino más bien el de una *definición* de la fragilidad; y que, por consiguiente, la fuerza explicativa del razonamiento sigue residiendo en la atribución de fragilidad a una hoja de ventana particular, es decir, en el enunciado de tipo legal L_1, no en una ley general acerca de todos los objetos frágiles.

Esa objeción tiene cierto peso cuando una característica disposicional representa exactamente un tipo de conducta como el legal: romperse por un impacto específico. Pero una característica disposicional, llamémosla *M*, del tipo aducido con fines explicativos por lo común puede manifestarse en toda una variedad de modos sintomáticos, según las circunstancias.[179] Por ejemplo, la imantación de una barra de hierro puede manifestarse en el hecho de que las limaduras de hierro se adhieren a sus extremos; pero también en el hecho de que uno de sus extremos atraerá al polo norte de una aguja magnética, y el otro extremo a su polo sur; y también en el hecho de que si esa barra se divide en dos, cada una de las partes presentará los dos tipos de disposiciones ya descriptos para la totalidad de la barra. Muchos de los "enunciados de síntomas" que caracterizan, así, alguna manera peculiar de manifestarse de *M* podrían considerarse como si expresaran una condición necesaria o suficiente para la presencia de *M*, y ésta misma como una característica ampliamente disposicional. Con respecto a tales características, la objeción mencionada no se aplica, como trataré de demostrar ahora.

Las oraciones de síntomas que expresan condiciones necesarias para *M* podrían adoptar la siguiente forma:

(9.2a) Si un objeto o individuo x tiene la propiedad M, entonces en condiciones de prueba o de estímulo del tipo S_1, x responderá regularmente de la manera R_1; en las condiciones S_2 de la manera R_2; etcétera.

Las oraciones de síntomas que expresan condiciones suficientes para *M* podrían adoptar la forma:

(9.2b) Si x se encuentra en condiciones del tipo S^1, entonces, si x responde de la manera R^1, tiene la propiedad M; si x se encuentra en las

[179] Ryle (1949, págs. 43-44) ha señalado que la atribución de una disposición habitualmente implica muchas proposiciones hipotéticas. Ya antes, Carnap había realizado un estudio formal mucho más completo de tales conceptos ampliamente disposicionales, en su ensayo *Testability and Meaning* [Testabilidad y significado] (1936-1937, Parte 2) en el que prevé específicamente la posibilidad de introducir un término científico por medio de un conjunto de oraciones de reducción, cada una de las cuales es una oración de síntoma en nuestro sentido. Para un examen que aclara aun más los temas aquí considerados, véase también Carnap (1956).

condiciones S^2, entonces, si x responde de la manera R^2, tiene la la propiedad de M; etcétera.[180]

Puede considerarse que cada oración de síntomas de cualquier tipo expresa un criterio parcial para la aplicación del término "M".

Concebir los enunciados de síntomas como si expresaran condiciones estrictamente necesarias o suficientes para M es, en muchos casos, una simplificación excesiva. Por ejemplo, en los enunciados de síntomas de la medicina y en la formulación de criterios parciales referentes a rasgos de carácter, creencias, deseos, etc., la relación entre M y sus manifestaciones sintomáticas a menudo deberá interpretarse como de carácter probabilístico. En este caso, las oraciones de síntomas podrían adoptar las siguientes formas estadísticas, correspondientes a (9.2a) y (9.2b):

(9.3a) Para objetos o individuos que se hallan en condiciones de prueba del tipo S_1 ($S_2, \ldots$), la probabilidad estadística de responder de la manera R_1 ($R_2 \ldots$) es r_1 ($r_2, \ldots$).

(9.3b) Para objetos o individuos que se hallan en condiciones de prueba del tipo S^1 ($S^2, \ldots$) y responden de la manera R^1 ($R^2, \ldots$), la probabilidad estadística de poseer la propiedad M es r' ($r'', \ldots$).

Sin embargo, para concentrarnos mejor en los problemas básicos que estamos examinando, limitaremos nuestra atención, por el momento, a rasgos ampliamente disposicionales, M, caracterizados por oraciones no probabilísticas de síntomas de las formas (9.2a) y (9.2b).

Sea U el conjunto de todas las oraciones de síntomas para M. Este conjunto evidentemente implica una oración, expresable en términos de "R_1", "S_1", "R_2", "S_2", ..., "R^1", "S^1", "R^2", "S^2", ... según la cual todo x que satisfaga algunas de las condiciones suficientes para M especificadas en U también satisfará cualquiera de las condiciones necesarias para M especificadas en U.[181] Como veremos en seguida, este enunciado normalmente tiene el carácter de una ley empírica general: y si las oraciones de síntomas para M tienen en conjunto implicaciones empíricas, no puede sostenerse que todas ellas son verdaderas simplemente por definición.[182]

Con el objeto de referirnos a un ejemplo anterior, una de las condiciones necesarias para que una barra de hierro sea magnética podría ser:

(9.4a) Si una barra de hierro x está imantada, entonces, si se colocan limaduras de hierro cerca de x (condición S_1), las limaduras se adherirán a sus extremos (respuesta R_1).

Y una de las condiciones suficientes podría ser:

[180] Los dos tipos de oraciones de síntoma, o criterios parciales de aplicación, aquí considerados, corresponden a los dos tipos básicos de oraciones de reducción" del estudio de Carnap (1936-1937); véase en particular la sección 8.

[181] Este enunciado es equivalente a lo que Carnap llama la "oración representativa" del conjunto U de oraciones de reducción para M; pues "representa, por así decir, al contenido fáctico" de U. Véase Carnap (1936-1937, págs. 451).

[182] Este punto está fundamentado e ilustrado con claridad, con referencia al concepto ampliamente disposicional de una persona *que desea* determinado estado de cosas, en Brandt y Kim (1963, págs. 428-429).

(9.4b) Si una barra de hierro x se encuentra en la vecindad de una aguja magnética (condición S^1), entonces, si uno de los extremos atrae al polo norte de la aguja y rechaza el polo sur, mientras que el otro extremo manifiesta la conducta opuesta (respuesta R^1), x está imantada (tiene la propiedad M).

Pero conjuntamente, estas dos oraciones de síntomas implican el enunciado general según el cual toda barra de hierro que satisfaga la condición de la aguja magnética también satisfará la condición referente a las limaduras de hierro: y este enunciado, sin duda, no es una verdad definicional, sino un enunciado que tiene el carácter de una ley empírica.

Así, por lo general, el conjunto U de enunciados de síntomas concernientes a un término ampliamente disposicional tiene consecuencias empíricas. Pero entonces es totalmente arbitrario concebir algunos de estos enunciados de síntomas como analítico-definicionales y asignar a otros el carácter de leyes empíricas; [183] pues esto sería equivalente a decretar que los primeros no están sujetos a modificaciones en caso de que se hallen elementos de juicio empíricos que entren en conflicto con las implicadas por el conjunto U. Ahora bien, en las ciencias empíricas, solamente las verdades lógicas y matemáticas gozan de tal inmunidad sin restricciones. Por consiguiente, es más adecuado considerar al conjunto total de enunciados de síntomas como parte del sistema de leyes generales que gobiernan al concepto en cuestión.

Supongamos ahora que con el fin de explicar por qué un objeto o individuo particular i se comportó de cierta manera R_3, se señala que i estuvo en una situación del tipo S_3 y que presenta una propiedad ampliamente disposicional M cuya presencia se caracteriza por la disposición a responder a S_1 de la manera R_1, a S_2 de la manera R_2, a S_3 de la manera R_3, etc. Este razonamiento explicativo puede esquematizarse así:

(C^1) i estuvo en una situación del tipo S_3
(C^2) i tiene la propiedad M
(9.5) (L) Todo x con la propiedad M, en una situación del tipo S_3, se comportará de la manera R_3.

(E) i se comportó de la manera R_3

Es claro que esta explicación tiene forma deductivo-nomológica; pues al enunciado general L, como acabamos de señalar, debe otorgársele el carácter de una ley empírica, no el de una mera "definición".

Pero las anteriores líneas sobre la "explicación disposicional" exigen una aclaración adicional. Lo dicho hasta ahora sugiere, por ejemplo, que atribuir a una barra de hierro la "propiedad ampliamente disposicional" de ser magnética equivale a atribuirle un conjunto de disposiciones simples, cada una de ellas caracterizada, en el sentido que reflejan nuestros enunciados de síntomas, por la asociación de algún tipo específico de "respuesta" manifiesta con ciertas "condiciones-estímulo" manifiestas. Esta

[183] Este punto ha sido tratado también en la Segunda Parte, *Concepciones de significación cognoscitiva.*

sería, sin embargo, una concepción demasiado simple. En efecto, los enunciados físicos generales concernientes a la propiedad de ser magnético incluyen, además de tales enunciados de síntomas, ciertas leyes generales que no representan tendencias disposicionales, pero que no son menos característicos del concepto de ser magnético que los enunciados de síntomas correspondientes. Entre ellas se encuentra la ley de que un campo magnético en movimiento producirá un campo eléctrico, lo cual implica que en un círculo de alambre cercano a un imán en movimiento se inducirá una corriente eléctrica, y lo cual —a su vez— implica un enunciado general concerniente a la respuesta dada por un amperímetro colocado en un círculo de alambre cercano a un imán en movimiento. Este último enunciado puede considerarse como otro enunciado de síntoma atinente a la propiedad de ser magnético, pero debe observarse que el síntoma especificado está asociado con la propiedad de ser magnético en virtud de principios teóricos que vinculan la característica dada con otros conceptos teóricos, tales como los de campo eléctrico, campo magnético y sus interrelaciones. Así, cuando un concepto como el de imán figura en una teoría, al aplicarlo a un objeto particular no estamos simplemente atribuyendo a este objeto un conjunto, por amplio que sea, de disposiciones a manifestar ciertos tipos de respuestas observables en condiciones de estímulo dadas y observables, sino que dicha asignación tiene también diversas implicaciones teóricas, incluyendo la atribución de otras características "ampliamente disposicionales".

Estas observaciones concernientes a los aspectos teóricos de conceptos ampliamente disposicionales son también atinentes, como se verá, al análisis del papel explicativo de las razones motivantes, que constituye el tema de la sección siguiente.

10. *El concepto de racionalidad y la lógica de la explicación por razones*

10.1. *Dos aspectos del concepto de racionalidad*

En esta sección me propongo examinar la lógica del método familiar de explicar las decisiones y acciones humanas en función de razones motivantes, método que suele considerarse muy diferente de los procedimientos explicativos de las ciencias naturales y que no permite el análisis por medio de modelos de leyes inclusivas.

En una explicación por razones motivantes la idea de racionalidad habitualmente desempeña un papel importante. Comenzaré, pues, con algunas observaciones sobre este concepto. Calificar determinada acción de racional equivale a proponer una *hipótesis empírica* y una *una apreciación crítica.* La hipótesis tiende a señalar que la acción se realizó por ciertas razones, a *explicar* aquélla como motivada por éstas. Esas razones incluirán los fines que presumiblemente persigue el agente y las creencias de que éste quizás abrigue en lo concerniente a la disponibilidad, propiedad y probable efectividad de los medios alternativos para alcanzar tales fines. La apreciación crítica implicada por la atribución de racionalidad afir-

maría que, juzgada a la luz de las creencias del agente, la acción que decidió realizar constituye una elección *razonable* o *adecuada* de los medios para lograr sus fines. Ahora examinaremos por turno estos dos aspectos del concepto de acción racional.

10.2. *La racionalidad como concepto normativo-crítico*

La clarificación de la idea crítica, o normativa, de acción racional exige la formulación de criterios claros de racionalidad que nos suministran normas para apreciar la racionalidad de acciones particulares, y que por ende pueden también guiarnos para adoptar decisiones racionales.

La racionalidad, en este sentido, es obviamente un concepto relativo. Que una determinada acción —o la decisión de llevarla a cabo— sea racional dependerá de los objetivos que pretenda alcanzar la acción y de la información empírica atinente al caso disponible en el momento de la decisión. Hablando en términos amplios, una acción racional si, sobre la base de la información dada, ofrece perspectivas óptimas de lograr sus objetivos. Consideremos ahora más detenidamente los conceptos fundamentales aducidos en esta caracterización: los conceptos de bases de información, objetivos de una acción y, finalmente, de racionalidad referida a bases y objetivos dados.

Para adoptar un curso racional de acción en la prosecución de determinados fines, tendremos que tomar en consideración toda la información disponible concerniente a cuestiones tales como las circunstancias particulares en las cuales debe emprenderse la acción; los diferentes medios por los cuales, en esas circunstancias, pueden alcanzarse los fines; y los efectos derivados y posteriores que puedan esperarse del uso de los diferentes medios disponibles.

La información empírica total disponible acerca de una decisión determinada puede concebirse como representada por un conjunto de oraciones, al que llamaré la *base informativa* de la decisión o la acción correspondiente. Esta manera de concebir la base empírica de una decisión toma en cuenta un hecho obvio pero importante: para juzgar la racionalidad de una decisión debe considerar no cuáles hechos empíricos —hechos particulares tanto como leyes generales— son atinentes en realidad al éxito o fracaso de la acción que se ha decidido realizar, sino de qué información concerniente a tales hechos dispone quien adopta la decisión. En verdad, una decisión puede calificarse de racional aunque se base en suposiciones empíricas incompletas o falsas. Por ejemplo, el historiador, precisamente para presentar como racional la acción de un personaje histórico, a menudo tendrá que suponer —y quizás hasta pueda demostrar sobre la base de elementos de juicio independientes— que el agente se hallaba incompletamente informado o que abrigaba creencias falsas respecto de cuestiones empíricas atinentes al caso.

Pero si bien no es necesario que sea verdadera la base informativa de una acción para que ésta sea racional, ¿no debe haber al menos buenas razones para creer que es verdadera? ¿No debe satisfacer dicha base los

requisitos de un sustento adecuado de elementos de juicio? Algunos autores consideran que ésta es una condición necesaria para que la acción sea racional, y esta opinión es muy plausible. Por ejemplo, como señala uno de sus defensores, Quentin Gibson: "Así, si alguien eludiera cuidadosa y deliberadamente una escalera porque creyera, sin fundamento, que caminar debajo de ella le traería mala suerte, no vacilaríamos en decir que esa persona actúa irracionalmente".[184]

Sin duda, a menudo entendemos la racionalidad en este sentido restringido. Pero si deseamos elaborar un concepto de acción racional que sea útil para explicar ciertos tipos de conducta humana, entonces parece preferible no imponer el requisito de que tenga el apoyo de elementos de juicio. En efecto, para explicar una acción en términos de las razones del agente, necesitamos saber lo que éste creía, pero no los fundamentos de tal creencia; y puede decirse que la persona actuó razonablemente, dada sus creencias.

Pasaré ahora de la base informativa de una decisión a sus objetivos. En casos muy simples puede concebirse una acción como tendiente a crear un particular estado de cosas, al que llamaré el estado anhelado. Pero aun en tales casos simples, algunos de los cursos de acción que, según la base informativa, están disponibles y probablemente produzcan el estado anhelado, pueden ser descartados porque violan ciertos principios restrictivos generales, tales como normas morales o legales, compromisos contractuales, convencionalismos sociales, las reglas del juego, etc. Por consiguiente, la acción proyectada tenderá a lograr el estado anhelado sin tal violación. Lo que llamaré su *objetivo total*, pues, puede caracterizarse por un conjunto E de oraciones que describan el estado anhelado y un conjunto N de normas restrictivas.

Nuevamente, como en el caso de la base empírica, no impondré el requisito de que haya "buenas razones" para adoptar los fines y las normas dadas: la racionalidad de una acción será entendida en un sentido estrictamente relativo, es decir, como su adecuación, juzgada sobre la base de la información dada, al logro del objetivo especificado.

¿Cómo puede definirse tal adecuación? Para las situaciones del tipo simple considerado y en las que es menester tomar una decisión, puede darse fácilmente una caracterización: si la base informativa contiene leyes generales en virtud de las cuales algunos de los cursos de acción disponibles permitiría lograr el objetivo total, entonces, evidentemente, cualquiera de esas acciones será racional en el contexto dado. Si la base informativa no permite distinguir ningún curso de acción disponible como medio suficiente para lograr el objetivo, puede, sin embargo, asignar una probabilidad numérica de éxito a cada uno de los diferentes cursos de acción disponibles; en este caso, será racional toda acción cuya probabilidad de éxito no sea inferior a la de cualquier alternativa disponible.

En muchos problemas de decisión racional, sin embargo, no es posible

[184] Gibson (1960, pág. 43). Los capítulos 4 y 14 de la obra de Gibson contienen muchas observaciones esclarecedoras sobre las cuestiones examinadas en esta sección.

concebir de este modo simple la información disponible, los objetivos y los criterios de racionalidad. En particular, nuestro enfoque es inaplicable cuando el objetivo de una acción propuesta no consiste en alcanzar un estado anhelado específico. Con frecuencia tal es el caso, como veremos ahora.

Para comenzar, aun cuando se tienda a un estado anhelado particular, la información disponible a menudo indicará que hay varios caminos posibles para alcanzarlo de modo definido o probable, cada uno de los cuales tiene un conjunto diferente de efectos derivados y consecuencias que no forman parte de él. Algunas de estas consecuencias incidentales se considerarán como más o menos deseables, otras como indeseables. En un modelo teórico de tales situaciones decisionales debe indicarse, pues, el objetivo total no simplemente describiendo el estado final deseado, sino también especificando la deseabilidad relativa de los diferentes resultados totales que puedan derivar de los cursos de acción disponibles.

En la teoría matemática de las decisiones se han construido varios modelos de elección racional en los que se supone que esas *deseabilidades* están especificadas en términos numéricos, como las llamadas utilidades de los diferentes resultados totales.

El caso en que la base informativa también especifica las "probabilidades" [185] de los diferentes resultados recibe el nombre de *decisión con riesgo*. Un criterio de racionalidad que ha ganado mucha aceptación para este caso es el de la *utilidad esperada maximizante*. La utilidad esperada que, sobre la base de la información dada, se halla asociada a un posible curso de acción se determina multiplicando para cada resultado posible de la acción, su probabilidad por su utilidad, y luego sumando los productos. Una acción, o la decisión de llevarla a cabo pues, es racional si su utilidad esperada es máxima, en el sentido de que no es inferior a la utilidad esperada de cualquier acción alternativa.

Otro problema de decisión que ha sido objeto de estudio matemático y que presenta considerable interés filosófico es el de la *decisión con incertidumbre*. En este caso se supone que la base informativa indica los diferentes cursos de acción disponibles y especifica, para cada uno de ellos, un conjunto de resultados posibles mutuamente excluyentes y conjuntamente exhaustivos, pero sin asignarles probabilidades; [186] finalmente, se

[185] Las probabilidades y utilidades mencionadas están sujetas a ciertos requisitos matemáticos que no podemos examinar dentro del contexto de este artículo. Su formulación clásica es la que le dieron von Neumann y Morgenstern (1947); claras exposiciones de sus requisitos y de las razones subyacentes a ellos se encontrarán en Luce y Raiffa (1957, caps. 1-4) y en Baumol (1961, caps. 17 y 18). Entre las cuestiones que aquí se pasan por alto se cuenta la muy importante que concierne a cómo debe entenderse el concepto de la probabilidad de los resultados dentro del contexto de la teoría de la decisión. Para una clase numerosa de problemas bastará prácticamente la familiar concepción estadística de la probabilidad como una frecuencia relativa a la larga, y sobre ella se basa en gran medida la actual teoría matemática de juegos y decisiones. Pero también se han propuesto otras concepciones, entre ellas la de la probabilidad inductiva o lógica de Carnap (1950 y 1962) y la de la probabilidad personal (véase Savage, 1954, en particular caps. 3 y 4).

[186] Hablando estrictamente, esta situación no puede darse en una teoría de lógica inductiva, como la de Carnap, para la cual la información empírica dada, sea la

supone que se ha asignado una utilidad a cada uno de los resultados posibles. A manera de ilustración, supongamos que se le regala a usted la bolilla de metal que saca en una única extracción realizada, a su elección, de una de dos urnas. Se le da a usted la información de que las bolillas de metal son todas del mismo tamaño; que la primera urna contiene bolillas de platino y bolillas de plomo en una proporción no especificada, y la segunda urna bolillas de oro y plata también en proporicones no especificadas. Supongamos que las utilidades que usted asigna al platino, el oro, la plata y el plomo están en la razón 1000:100:10:1. ¿De cuál de las urnas sería racional hacer la extracción? En la reciente teoría de la decisión, se han propuesto varios criterios totalmente diferentes de elección racional con incertidumbre. Quzás el más conocido de ellos sea la *regla del maximin.* Ella prescribe maximizar la utilidad mínima, es decir, elegir un curso de acción cuyo peor resultado posible sea el menos tan bueno como el peor resultado posible de cualquier alternativa. En nuestro ejemplo, esto exige realizar una extracción de la segunda urna, pues en el peor de los casos obtendrá usted una bolilla de plata, mientras que el peor resultado de una extracción de la primera urna le daría una bolilla de plomo. Esta regla representa, claramente, una política de extrema cautela que refleja la máxima pesimista: actúa en la suposición de que tu acción tendrá el peor resultado posible.

Una política alternativa, expresada por la *regla del maximax,* refleja la esperanza optimista de que nuestra acción conducirá al mejor resultado posible. Ella nos prescribe elegir un curso de acción cuyo mejor resultado posible sea al menos tan bueno como el mejor resultado posible de cualquier acción alternativa que podamos seguir. En nuestro ejemplo, la decisión adecuada, según esta regla, sería extraer de la primera urna, pues en el mejor de los casos obtendremos una bolilla de platino, mientras que una extracción de la segunda urna nos dará en el mejor de los casos una bolilla de oro.

Se han propuesto diversas reglas alternativas interesantes para el caso de la decisión con incertidumbre, pero no es necesario para nuestros fines considerarlas aquí.[187]

Los modelos matemáticos que hemos caracterizado brevemente no brindan mucha ayuda para dar una solución racional a los graves y complejos problemas de decisión que debemos abordar en la vida cotidiana. En efecto, en estos casos, estamos lejos, por lo común, de tener los datos requeridos por nuestros modelos: a menudo no tenemos idea clara de los cursos de acción disponibles, ni podemos especificar los resultados posibles, para no hablar ya de sus probabilidades y utilidades. Sin embargo, en contextos en los que se dispone de tal información, se ha aplicado con todo éxito la teoría matemática de la decisión aun a problemas bastante complicados, por ejemplo, al control industrial de calidad y a algunas fases de la planificación estratégica.

---

que fuere, siempre asigna una probabilidad lógica definida a cada uno de los enunciados que describen uno de los resultados posibles.

[187] Puede encontrarse una exposición de estas reglas, por ejemplo, en Luce y Raiffa (1957, cap. 13) y en Baumol (cap. 19).

Pero sea cual fuere su valor práctico, estos modelos contribuyen, creo, a realizar el esclarecimiento analítico del concepto de acción racional. En particular, ponen de relieve el carácter complejo y múltiple relativo de este concepto; y muestran que algunas de las caracterizaciones de la acción racional expuestas en la literatura filosófica son de una claridad y simplicidad engañosa. Por ejemplo, Gibson, en su minucioso y claro estudio, observa: "Hay varias maneras alternativas de lograr un fin. Actuar racionalmente ... es elegir lo que, según los elementos de juicio, es la *mejor manera* de alcanzarlo";[188] y se refiere a "un punto lógico elemental, a saber, que dados ciertos elementos de juicio, sólo puede haber una solución correcta al problema de determinar cuál es la mejor manera de lograr cierto fin".[189] Gibson no ofrece ningún criterio para saber cuál es la mejor solución; pero, sin duda, lo que afirma aquí no es un punto lógico elemental, y en realidad no es verdadero. En efecto, en primer lugar, aun cuando la situación decisional sea de un tipo con respecto al cual pueda suponerse que se dispone de un criterio convenido de elección racional —por ejemplo, el principio de las utilidades maximizantes esperadas—, tal criterio puede calificar a varios cursos diferentes como igualmente racionales. En segundo lugar, y más importante aún, hay varios tipos de decisiones, como la decisión con incertidumbre, con respecto a los cuales no hay ni siquiera acuerdo acerca de un criterio de racionalidad, y el criterio del maximin se opone al de maximax y ambos, a diversas reglas alternativas.

Es importante tener presente que los diferentes criterios rivales de racionalidad no reflejan diferencias en la evaluación de los diversos fines alcanzables sobre la base de la información dada: todas las reglas rivales a las que aludimos presuponen que se han determinado de antemano las utilidades de esos fines. Las diferentes reglas o criterios de racionalidad para la decisión reflejan más bien diferentes actitudes inductivas y en algunos casos, como vimos, diferentes grados de optimismo o pesimismo en cuanto a lo que cabe esperar del mundo, y por ende diferentes grados de audacia o cautela en la elección de un curso de acción.

La diversidad de reglas conflictuales propuestas para la decisión con incertidumbre sugiere el problema de si no sería posible especificar algún sentido único de racionalidad que sea independiente de tales diferencias de concepción y del cual pueda mostrarse que es más adecuado que las concepciones de la racionalidad reflejadas por los criterios rivales que hemos mencionado. Las perspectivas de especificar tal sentido son oscuras, en verdad, como lo indican —nuevamente— algunos resultados de la teoría matemática de la decisión. Específicamente, es posible formular un conjunto de desiderátums generales, o condiciones de adecuación, para toda regla de decisión propuesta, y mostrar que si bien cada uno de los desiderátums parece perfectamente razonable y —por así decir— "esencial" para la elección racional, sin embargo, (*a*) toda regla de decisión que se ha propuesto en la bibliografía sobre el tema viola uno o más de los desiderátums y que, en realidad, (*b*) a pesar de su plausibilidad intuitiva, los desiderá-

188 Gibson (1960); bastardillas del autor.
189 Gibson (1960, pág. 162).

tums son lógicamente incompatibles.[190] Este resultado debe servir de advertencia en contra de la suposición de que la idea de racionalidad, o de la mejor manera de actuar en una situación dada, es razonablemente clara y de que la formulación de criterios que haga explícita la noción es una tarea explicativa básicamente trivial y, quizá, tediosa.

Las consideraciones esbozadas sobre la noción crítica o normativa de racionalidad tienen importantes implicaciones para el uso explicativo de la idea de acción racional, como veremos a continuación.

10.3. *La racionalidad como concepto explicativo*

A menudo se explican las acciones humanas en función de razones motivantes. Las consideraciones precedentes sugieren que una enunciación completa de esas razones tendrá que indicar los objetivos del agente, así como sus credenciales acerca de los medios de que dispone y de sus probables consecuencias. Y la explicación tenderá a mostrar que la acción era de esperar, en vista de esos objetivos y esas creencias. Tales explicaciones se basan, pues, para emplear las palabras de Peters, en la "suposición oculta" de que "*los hombres son racionales* porque adoptan medios que conducen a fines, si disponen de información y desean esos fines".[191] El concepto de racionalidad, pues, se utiliza como hipótesis explicativa. Examinemos ahora la lógica de tales explicaciones.

10.3.1. *El concepto de Dray de explicación racional.* Elijamos como punto de partida el estimulante y sugerente estudio de Dray sobre tales explicaciones y, en particular, de su papel en la investigación histórica,[192] estudio que lo llevó a la conclusión de que "la explicación de la conducta humana individual tal como se da habitualmente en la historia tiene caracteres que hacen peculiarmente ineficaz el modelo de leyes inclusivas".[193] Dray se refiere al tipo de explicación al que hemos aludido, es decir, a la explicación por razones motivantes, como a una *explicación racional,* porque, dice, "expone la *justificación* de lo hecho" ofreciendo "una reconstrucción del *cálculo,* realizado por el agente, de los medios que se deben adoptar para alcanzar su fin elegido, a la luz de las circunstancias en las que se encuentra. Para explicar la acción necesitamos saber qué consideraciones lo convencieron de que debía actuar como lo hizo".[194] Pero Dray atribuye a la explicación racional otra característica que asigna claramente un papel esencial al concepto evaluativo o crítico de racionalidad. Según él, el "objetivo de tal explicación es mostrar que lo que se hizo era lo que se debía hacer por las razones dadas, y no sólo lo que se hace en tales

[190] Véase detalles en Luce y Raiffa (1957, cap. 13, en particular las secciones 3 y 4).

[191] Peters (1958, pág. 4); las bastardillas son mías. Otra declaración concerniente al uso explicativo y predictivo de la suposición de racionalidad se encontrará en Gibson (1960, pág. 164).

[192] Véase en particular Dray (1957, cap. 5) y Dray (1963).

[193] Dray (1957, pág. 118).

[194] Dray (1957, págs. 122 y 124); bastardillas del autor.

ocasiones, quizá de acuerdo con ciertas leyes".[195] Luego, "las razones aducidas, para que sean explicativas de una manera racional, deben ser *buenas* razones, al menos en el sentido de que *si* la situación hubiera sido tal como el agente la veía ... entonces, lo que se hizo fue lo que se debía hacer".[196] Para demostrar que el agente tenía buenas razones para emprender su acción, una explicación racional debe, pues, invocar no una ley empírica general, sino un "*principio de acción*" que exprese un juicio de la forma: "En una situación de tipo $C_1 \ldots, C_n$, lo que se debe hacer es x".[197] Así, tales explicaciones contienen "un elemento de *apreciación* de lo que se ha hecho". [198] Precisamente, en esta referencia a un principio de acción que expresa una norma de adecuación o racionalidad es donde Dray ve la diferencia esencial entre las explicaciones racionales y aquellas que explican un fenómeno subsumiéndolo bajo leyes generales inclusivas que describen ciertas uniformidades, pero no hacen apreciaciones.

Dray no especifica más el carácter de las "situaciones" a las que alude en sus principios de acción pero para hacer justicia a su propósito, debe considerarse que tales situaciones incluyen puntos tales como: (a) el fin que el agente quería alcanzar; (b) las creencias del agente concernientes a las circunstancias empíricas en las que tuvo que actuar y a los medios de que disponía para el logro de su objetivo; (c) normas morales, religiosas y de otro tipo, a las que el agente adhería. En efecto, sólo cuando se especifican estos puntos tiene sentido plantear la adecuación de lo que el agente hizo en la situación dada.

Parece correcto, pues, decir que según la concepción de Dray una explicación racional responde a una pregunta de la forma "¿por qué el agente A hizo X?", ofreciendo un explanans del siguiente tipo (en lugar de escribir "$C_1 \ldots C_n$, como hace Dray, escribimos "C" para mayor brevedad, pero debemos recordar que la situación aludida puede ser muy compleja):

A estaba en una situación de tipo C
En una situación de tipo C, lo que se debe hacer es X.

Pero este enfoque de la explicación racional presupone un criterio de racionalidad que, para el tipo dado de situación, destaque un particular curso de acción como *lo* que se debe hacer; y ya vimos antes que esta presuposición es muy discutible.

Pero lo más importante es que aun cuando se dispusiera de tal criterio, una explicación de la forma considerada no podría explicar por qué A hizo X. En efecto, según el requisito de adecuación expuesto en la sección 2.4 de este capítulo, toda respuesta adecuada a la pregunta de por qué se produjo un suceso determinado tendrá que suministrar una información que, si se la acepta como verdadera, brindará buenas razones para creer que el hecho se produjo. Ahora bien, la información de que el agente A se encontraba en una situación de tipo C y de que, en tal situación, la

195 Dray (1957, pág. 124).
196 Dray (1957, pág. 126); bastardillas del autor.
197 Dray (1957, pág. 132); bastardillas del autor.
198 Dray (1957, pág. 124); bastardillas del autor.

acción racional que se debe emprender es *x*, ofrece fundamento para creer que *A habría actuado racionalmente si hubiera hecho x*, pero no para creer que A hizo x, de hecho.[199] Para justificar esta última creencia es claro que necesitamos otra suposición explicativa, a saber que —al menos en el momento en cuestión— *A* era un *agente racional* y, por ende, estaba *dispuesto* a hacer todo lo que fuera racional en esas circunstancias.

Pero cuando se agrega esta última suposición la respuesta a la pregunta "¿por qué *A* hizo *x*?" adopta la siguiente forma:

A estaba en una situación de tipo *C*

A era un agente racional

(Esquema *R*) En una situación de tipo *C*, todo agente racional hará *x*

Por lo tanto, *A* hizo *x*

Este esquema de la explicación racional difiere en dos aspectos de lo que considero el enfoque de Dray: primero, se agrega explícitamente la suposición de que *A* era un agente racional; y segundo, se reemplaza el principio evaluativo o estimulativo de acción, que especifica lo que debe hacerse en la situación *C*, por una generalización empírica que enuncia cómo actuarán los agentes racionales en situaciones de ese tipo. De este modo, el enfoque de Dray falla justamente en el punto en que se propone revelar una diferencia lógica entre las explicaciones que aluden a razones subyacentes y las que realizan una subsunción bajo leyes generales, pues con el fin de asegurar la eficacia explicativa de una explicación racional vimos que era menester reemplazar el principio de acción normativo de Dray por un enunciado que tiene el carácter de una ley general. Péro restituye a la explicación la forma de ley inclusiva.

El hecho de que la función estimativa que Dray considera esencial para la explicación racional no tiene alcance explicativo se demuestra también por la siguiente consideración: Las dudas concernientes a una explicación dada en términos de una justificación especificada no pueden expresarse con sentido en la forma "¿era *X* realmente lo que se debía hacer en esas circunstancias?", pero podrían adoptar la forma "¿estaba *A* realmente inclinado a considerar a *X* como lo que se debía hacer?" Por consiguiente, sería inútil argüir, en defensa de una explicación propuesta, que *X* era en realidad (por alguna norma teórica de racionalidad) "lo que se debía hacer", mientras que sí sería eficaz mostrar que *A* se hallaba en general dispuesto a hacer *X* en las circunstancias del tipo especificado. Y el alcance explicativo de esta última información sería totalmente independiente de que la acción considerada se ajustara o no a las normas de racionalidad de quien da la explicación o hace la pregunta.

[199] La misma objeción, en efecto, ha planteado Passmore en el siguiente comentario sobre la concepción de Dray: "...la explicación por referencia a un 'principio de acción' o a una 'buena razón' no es, en sí misma ninguna explicación ... Pues una razón puede ser una 'buena razón', en el sentido de ser un principio al cual *podríamos* apelar como justificación de la propia acción, sin que de hecho haya ejercido la menor influencia sobre nosotros". Passmore (1958, pág. 275); bastardillas del autor.

Al discrepar con el análisis que hace Dray de la explicación racional, no pretendo negar que una explicación realizada en términos de razones motivantes pueda tener resonancias evaluativas: lo que sostengo es solamente que el hecho de que una explicación contenga o sugiera una apreciación crítica es ajeno a su fuerza explicativa; y que una apreciación sola, por medio de lo que Dray llama un principio de acción, no puede explicar en absoluto por qué A, de hecho, hizo X.

10.3.2. *La explicación por razones de carácter ampliamente disposicional.* La noción de agente racional incluida en el esquema R debe tomarse, por supuesto, como un concepto psicológico regido por criterios objetivos de aplicación; toda connotación normativa o evaluativa que pueda tener carace de importancia para su uso explicativo. Sin duda, los preconceptos normativos acerca de cómo debería comportarse una persona verdaderamente racional pueden influir sobre la elección de los criterios descriptivos para determinar si un agente es racional, así como la elaboración de tests y, por ende, la selección de criterios objetivos para determinar la inteligencia, la aptitud verbal, la aptitud matemática, etc., sufrirán la influencia de concepciones y normas presistemáticas. Pero el uso descriptivo-psicológico del término "agente racional" (así como el de los términos "CI", "aptitud verbal, "aptitud matemática", etc.) debe estar regido, entonces, por las reglas empíricas objetivas de aplicación que se hayan adoptado, independientemente de si esta o aquella persona (por ejemplo la que propone una explicación racional o aquella a quien se dirige) consideren esas reglas objetivas como concordantes con sus propios patrones normativos de racionalidad.

Sean cuales fueren los criterios empíricos específicos por los que se la pueda caracterizar, la racionalidad —en el sentido descriptivo-psicológico— es una *característica ampliamente disposicional*; decir de alguien que es un agente racional es atribuirle, por implicación, un complejo conjunto de disposiciones. Cada una de estas puede concebirse como una tendencia a comportarse —de modo uniforme o con cierta probabilidad— de una manera característica en condiciones de un cierto tipo, cuya especificación completa quizá deba incluir información acerca de los objetivos y creencias del agente, acerca de otros aspectos del estado psicológico y biológico en el que se encuentra y acerca de su ambiente. Explicar una acción en función de las razones del agente y de su racionalidad es, pues, presentar la ación como conforme a esas tendencias generales o como una manifestación de ellas.[200]. Según que las oraciones que expresen las tendencias aludidas sean de forma estrictamente universal o de forma estadística, como (9.3a) o (9.3b), la explicación disposicional resultante será de carácter deductivo o inductivo-probabilístico. Pero en ambos casos subsumirá el caso particular dado bajo una uniformidad general. Pero debemos ahora ampliar esta

[200] Este enfoque, por supuesto, está básicamente de acuerdo con la concepción general presentada por Ryle (1949). Para una caracterización aclarada, de acuerdo con las ideas de Ryle, de la fuerza de las explicaciones que se refieren a los deseos, intenciones, planes, etc., de un agente, véase Gardiner (1952, Parte IV, sección 3); y también el examen expositivo y crítico de Dray (1957, págs. 144 y *passim*). |

breve caracterización general, así como también aclarar algunos puntos particulares.

Para comenzar, las disposiciones implicadas por el concepto psicológico de agente racional no son simplemente disposiciones a responder a estímulos externos especificables con ciertos modos característicos de conducta manifiesta. Difieren, a este respecto, de algunas, al menos de las disposiciones implicadas cuando decimos de una persona que es alérgica al polen; pues decir esto implica, entre otras cosas, que presentamos los síntomas de un resfrío, cuando estamos expuestos al polen. Al decir de alguien que es un agente racional, afirmamos por implicación que se comportará de manera característica si se encuentra en determinados tipos de situaciones. Pero no es posible describir tales situaciones simplemente en términos de condiciones ambientales y estímulos externos, pues ellas incluyen de modo típico los objetivos del agente y sus creencias atinentes a la cuestión. Para destacar esta diferencia, podríamos decir que las disposiciones implicadas al atribuir racionalidad a una persona son *disposiciones de orden superior*; pues las creencias y fines en respuesta a los cuales, por decir así, el agente actúa de una manera característica no son estímulos externos manifiestos, sino más bien características ampliamente disposicionales del agente. En verdad, atribuir a alguien una creencia o fin particular equivale a afirmar por implicación que en ciertas circunstancias tenderá a camportarse de ciertas maneras que indican o son síntomas de su creencia o fin.

Pero hay aún otra razón por la cual debemos evitar una concepción estrechamente disposicional de las creencias, objetivos y racionalidad de un agente, y la frase aclaratoria "ampliamente disposicional" tiende a recordar también este punto: un enunciado que atribuye a una persona ciertos objetivos o creencias, o la propiedad de ser un agente racional, implica un conjunto de otros enunciados que atribuyen a esa persona determinados cúmulos de disposiciones, *pero no es equivalente a dicho conjunto.*

Para aclarar y dar sustento a esta idea, apelaré primero a un caso análogo tomado de la física. Decir de un cuerpo que está eléctricamente cargado o que es magnético es atribuirle, *por implicación,* conjuntos de disposiciones a responder de maneras características o sintomáticas a diversos procesos de prueba. Pero esto no es todo lo que se afirma, pues los conceptos de carga eléctrica, magnetización, etc., están gobernados por una red de principios teóricos que interconectan un gran número de conceptos físicos. Conjuntamente, estos principios teóricos determinan un conjunto indefinidamente grande de consecuencias empíricamente testables, entre ellas diversos enunciados disposicionales que suministran criterios operacionales para establecer si un cuerpo está cargado eléctricamente, es magnético, etc. Así, las suposiciones teóricas subyacentes contribuyen de modo esencial a lo que se afirma mediante la atribución de esas propiedades físicas. En verdad, sólo en conjunción con tales suposiciones teóricas básicas un enunciado que atribuya una carga eléctrica a un cuerpo dado implica un conjunto de enunciados disposicionales; mientras que todo el conjunto de enunciados disposicionales no implica el enunciado acerca de la carga, y menos aún los principios teóricos básicos.

Ahora bien, sin duda los conceptos psicológicos que sirven para indicar las creencias, objetivos y normas, etc., de una persona no están incluidos en una red teórica tan vasta y explícita como la de la teoría electromagnética. Usamos, sin embargo, esos conceptos psicológicos de una manera que presupone claramente ciertas conexiones similares a las que podríamos llamar *conexiones cuasi-teóricas*.[201] Por ejemplo, suponemos que la conducta que manifiesta una persona que persigue cierto objetivo dependerá de sus creencias; y recíprocamente. Así, la atribución a Enrique de la creencia de que las calles están fangosas implica que se pondrá galochas sólo sobre la base de suposiciones adecuadas acerca de sus objetivos y también acerca de sus otras creencias,[202] por ejemplo, que quiere salir, quiere mantener sus pies secos, cree que sus galochas servirán a tal fin, no está demasiado apurado para ponérselas, etc. Está claro que esto refleja las suposiciones de muchas interdependencias complejas entre los conceptos psicológicos en cuestión; y son estas suposiciones las que determinan nuestras expectativas acerca de las manifestaciones conductales, incluso las acciones manifiestas, que tendrán un rasgo psicológico en un caso particular.

Rechazar la concepción de tales características como meros conjuntos de disposiciones conductales no significa volver a introducir en la máquina el espíritu, tan hábil y suitilmente exorcizado de Ryle y ya antes —de modo más resumido, pero sobre fundamentos esencialmente similares— por el conductismo lógico de Carnap.[203] La cuestión es, más bien, que para caracterizar los rasgos psicológicos en discusión, no sólo debemos considerar sus implicaciones disposicionales, que suministran criterios operacionales para atribuir ciertas creencias, objetivos, etc., sino que también debemos tomar en cuenta las suposiciones cuasi-teóricas que los vinculan; ya que también éstos gobiernan el uso de esos conceptos y no están implicados lógicamente en los conjuntos de enunciados disposicionales asociados.

10.3.3. *Interdependencia epistémica de las atribuciones de creencias y las atribuciones de fines.* Las conexiones cuasi-teóricas a las que hemos aludido plantean un problema que exige por lo menos una consideración breve. Para nuestros propósitos bastará examinar una de sus formas que es de fundamental importancia para la idea de explicación racional. ¿Qué tipos de disposiciones atribuimos a una persona, por implicación, cuando afirmamos que tiene ciertas creencias u objetivos específicos? El enunciado según el cual Enrique *quiere* un vaso de agua implica, entre otras cosas,

[201] Algunos principios cuasi-teóricos plausibles para el concepto de agente con un cierto objetivo o que "desea" un cierto estado de cosas se encontrarán en Brandt y Kim (1963, pág. 427), quienes sugieren que el concepto "desea" podría ser provechosamente considerado como un construcción teórica. Tolman (1951) presenta, en un esbozo algo esquemático y más bien programático, una teoría de la acción basada en un modelo psicológico que incluye, entre sus "variables intermedias", la "matriz de valores de creencia" y el "sistema de necesidades" del agente, pero que también toma en consideración, correctamente, las condiciones externas en las que se realiza la acción.

[202] Sobre este punto, véase Chisholm (1962, págs. 513 y sigs., en particular la pág. 517).

[203] Véase Ryle (1949), Carnap (1938) y, para una exposición más técnica, Carnap (1936-1937).

que está dispuesto a beber un líquido que se le ofrezca, siempre que *crea* que es agua potable (y siempre que no tenga razones opuestas para rechazarlo). De este modo, la atribución en este caso de un objetivo tiene implicaciones concernientes a la conducta manifiesta característica sólo cuando se la toma juntamente con la atribución de creencias adecuadas. De manera análoga en nuestro ejemplo anterior la hipótesis de que Enrique *cree* que las calles están fangosas implica la aparición de una conducta manifiesta característica sólo cuando se la toma junto con hipótesis adecuadas acerca de los *objetivos* de Enrique.

En verdad, parece que una hipótesis acerca de los objetivos de un agente puede considerarse, en general, como si implicara la aparición de una acción específica manifiesta sólo cuando se la asocia a hipótesis adecuadas acerca de sus creencias, y viceversa. Luego, hablando en términos estrictos, un examen de la conducta del agente puede servir para someter a prueba suposiciones acerca de sus creencias o acerca de sus objetivos, no por separado, sino sólo en pares adecuados. Es decir, las atribuciones de creencias y las atribuciones de fines son *epistémicamente interdependientes.*

Este hecho, sin embargo, no impide determinar las creencias o los objetivos de una persona, pues a menudo disponemos de una buena información anterior acerca de uno de los elementos interdependientes, con lo cual es posible someter a prueba una hipótesis respecto del otro elemento, estableciendo cómo actúa dicha persona en ciertas situaciones. Por ejemplo, si tenemos buenas razones para suponer que nuestra persona es honesta, que trata de "decir la verdad", entonces sus respuestas a nuestras preguntas pueden brindar una indicación confiable acerca de sus creencias. Recíprocamente, a menudo podemos someter a prueba una hipótesis acerca de los objetivos de una persona, examinando su conducta en ciertas situaciones críticas porque tenemos buenas razones para suponer que abriga determinadas creencias.

Pero la interdependencia epistémica a la que aludimos plantea la cuestión de saber si una explicación por razones motivantes exige la suposición explicativa de que la persona actuante fue, al menos en el momento considerado, un agente racional. Puede verse cómo surge este problema examinando con mayor detenimiento los criterios de prueba para las atribuciones de creencias y para las atribuciones de fines.

Supongamos que conocemos las creencias de un agente y queremos someter a prueba la hipótesis de que quiere lograr el fin *F*. ¿Qué tipo de acción, exactamente, implica esta hipótesis? El criterio usado en tales casos parece ser aproximadamente éste: si *A* quiere en realidad lograr *B*, entonces seguirá un curso de acción que, a la luz de sus creencias, le brinde las mayores probabilidades de éxito. En el lenguaje de nuestro examen anterior, pues, la prueba de nuestra atribución de fines parece presuponer la afirmación de que *A* elegirá un curso de acción que será racional con respecto a sus objetivos y creencias. Esto significaría que la manera como usamos las acciones de una persona como elemento de juicio para establecer sus fines ya contiene en sí misma la suposición de la racionalidad. Un comentario análogo puede hacerse de la manera como, normalmente, usa-

mos las acciones de una persona cuyos objetivos conocemos como elemento de juicio para discernir sus creencias.[204] Pero esto parece desacreditar la concepción según la cual la explicación racional supone, del modo sugerido en el esquema *R*, una hipótesis explicativa que afirme que la persona implicada era un agente racional. En efecto, las consideraciones esbozadas sugieren que esta hipótesis es siempre verdadera en virtud de una convención implícita en nuestros criterios de prueba para la atribución de objetivos y creencias motivantes al agente. Si esto es lo que sucede en general, entonces no es posible violar la suposición de la racionalidad. Toda violación aparente sólo sería tomada como indicio de que nuestras conjeturas acerca de las creencias del agente o de sus objetivos, o de ambos, eran equivocadas. En efecto, tal será nuestro veredicto.

Pero ¿será siempre así? Creo que hay diversos tipos de circunstancias con respecto a los cuales podemos conservar nuestras suposiciones acerca de las creencias y objetivos del agente y abandonar, en cambio, la suposición de la racionalidad. Ante todo, al adoptar una decisión respecto del curso de acción que seguirá, una persona puede pasar por alto ciertos elementos informativos atinentes al caso que ella sin duda considera verdaderos y que, de haber sido tomados adecuadamente en consideración, le habrían inspirado un curso de acción diferente. En segundo lugar, el agente puede pasar por alto ciertos aspectos del objetivo total que está tratando de conseguir, y puede decidir realizar una acción que no es racional, juzgada según sus objetivos y creencias. En tercer lugar, aunque el agente tomara en cuenta todos los aspectos de su objetivo total, así como la información atinente al caso de que dispone, y aun cuando realizara un deliberado "cálculo de los medios que se deben adoptar para lograr el fin elegido" (para repetir una cita anterior de Dray), el resultado puede no ser una decisión racional por alguna falla lógica en el cálculo. Evidentemente, puede haber sólidos elementos de juicio, en algunos casos, que revelen que un agente no actuó con racionalidad de alguna de las maneras indicadas; en realidad, si tomó su decisión bajo la presión del tiempo o en estado de tensión emocional, fatiga u otras influencias perturbadoras, tales desviaciones de la racionalidad se considerarían como muy probables. (Esto refleja otra de las conexiones cuasi-teóricas, entre los diversos conceptos psicológicos que desempeñan cierto papel en las explicaciones por razones o por motivos.)

Para resumir, pues, las convenciones implícitas en los criterios que gobiernan la atribución de fines y creencias a agentes humanos no garantizan la racionalidad de las acciones humanas, puede haber buenas razones para atribuir a un agente ciertos objetivos y creencias y, sin embargo, reconocer que su acción no era una exigencia racional de sus objetivos y creencias.

10.3.4. *La acción racional como concepto-modelo explicativo.* Para aclarar aun más el papel que la suposición de la racionalidad desempeña en las explicaciones por razones motivantes, puede resultar útil preguntarse si

[204] Véase, por ejemplo, el examen de Churchman (1961, págs. 288-291) que ilustra este punto.

no sería posible considerar el concepto de agente racional como un modelo explicativo idealizado comparable al concepto explicativo de gas ideal, es decir, de un gas que satisfaga exactamente las leyes de Boyle y Charles. Ningún gas real satisface exactamente estas leyes, pero hay una amplia gama de condiciones dentro de las cuales muchos gases se ajustan bastante estrechamente a la explicación que da el modelo de las interrelaciones entre temperatura, presión, y volumen. Además, hay leyes más generales, pero menos simples, como las de van der Waals, Clausius y otras que explican, en buena medida, las desviaciones del modelo ideal que presentan los gases reales.

Quizás el concepto de agente racional pueda considerarse de manera análoga como un modelo explicativo caracterizado por una "ley ideal", según la cual las acciones del agente son estrictamente racionales (en el sentido establecido por algún criterio específico) con respecto a sus objetivos y creencias. ¿Cómo podría aplicarse esta concepción propuesta? ¿Cómo podría caracterizarse con precisión un modelo explicativo de acción racional, y cómo se lo podría aplicar y someter a prueba?

Como observamos antes, el concepto de racionalidad no es en modo alguno tan claro e inequívoco como se supone a veces en la literatura sobre explicación racional. Pero supongamos que el uso explicativo propuesto del concepto se limite, para comenzar, a casos de un tipo relativamente simple y para los cuales pueda formularse e incorporarse a nuestro modelo algún criterio preciso de racionalidad. Quedaría en pie la cuestión de cómo aplicar el modelo a casos particulares, cómo establecer si una acción determinada se ajusta o no de hecho al criterio de racionalidad incluido en el modelo. Y esto plantea un problema desconcertante que no es el problema práctico de *establecer* las creencias y las acciones de un agente en un caso dado, sino el problema conceptual de saber qué debe *entenderse* por creencias y objetivos de un agente en un momento dado y por cuáles medios lógicos se los puede caracterizar adecuadamente. Permítaseme desarrollar con brevedad este punto.

Puede admitirse, sin duda, que una persona sostiene muchas creencias de las que no es consciente en el momento, pero que pueden hacerse explícitas por diversos medios. En realidad, puede sostenerse que una persona cree muchas cosas en las que nunca ha pensado y en las que quizá nunca llegue a pensar en toda su vida. Si cree que siete más cinco son doce, seguramente podemos considerar que cree también que siete gallinas batarazas y otras cinco sumarán doce gallinas batarazas, aunque nunca haya abrigado conscientemente esta creencia particular. En general, puede considerarse que un hombre cree ciertas cosas que son consecuencias de otras cosas que cree, pero sin duda no puede sostenerse que cree *todas* esas consecuencias, ya que —para mencionar sólo una razón— su perspicacia lógica es limitada.

Por consiguiente, si bien en un modelo teórico del concepto normativo o crítico de decisión racional la base informativa puede concebirse como un conjunto de enunciados cerrados con respecto a una adecuada relación de derivación lógica, no puede transferirse esta suposición a un modelo explicativo de la decisión racional. En particular, una persona puede dar

su aprobación a uno de los enunciados de un par de ellos lógicamente equivalentes y negársela al otro, aunque ambos expresen la misma proposición. Parece claro, pues, que no es posible concebir los objetos de las creencias de una persona como proposiciones, cada una de las cuales pueda representarse por un enunciado cualquiera de un conjunto de enunciados equivalentes: al especificar las creencias del agente, el modo de formulación es esencial. (Esta peculiaridad parece muy afín a la que Quine ha llamado la opacidad referencial de las oraciones de creencia.)[205]

Es de presumir, pues, que en una concepción de la acción racional basada en un modelo explicativo las creencias del agente deben ser representadas por algún conjunto de oraciones que no sea cerrado con respecto a la derivación lógica. Pero ¿cuál conjunto? Por ejemplo, ¿debe considerarse que el conjunto de creencias de una persona incluye todas las oraciones que puedan inducirlo a admitir mediante preguntas y razonamientos atinentes a cada caso, por muchas o complejas que sean? Es obvio que tal interpretación carece de base si estamos interesados en especificar un conjunto de creencias que se consideren como factores motivantes en la explicación de una acción realizada por el agente. Dónde trazar, conceptual no prácticamente, la línea demarcatoria del conjunto de creencias es una cuestión desconcertante y oscura.

Observaciones similares son aplicables al problema de cómo caracterizar los objetivos totales de un agente en una determinada situación de adopción de decisiones.

Por consiguiente, aunque en un modelo normativo-crítico de decisión siempre se juzga la racionalidad con referencia a la base informativa total y al objetivo total especificado, sería contraproducente incorporar a un modelo explicativo de acción racional el principio de que un agente racional actúa óptimamente, según criterios específicos, sobre la base del conjunto total de sus objetivos y creencias: simplemente, esta noción es demasiado oscura.

10.3.5. *El modelo de un agente conscientemente racional.* La observación de que muchas explicaciones presentan una acción como determinada por razones que, se presume, el agente tomó conscientemente en cuenta al adoptar su decisión, parece sugerir una solución. Digamos que una persona es un *agente conscientemente racional* (en un momento determinado) si (en ese momento) sus acciones son racionales (en el sentido de algún criterio claramente especificado) con respecto a aquellos de sus objetivos y creencias que toma conscientemente en cuenta para llegar a su decisión.

Para explorar la aplicabilidad potencial de este modelo de un agente conscientemente racional, consideraremos la redacción de Bismarck del llamado telegrama de Ems, que desempeñó un papel fundamental en el

[205] Véase Quine (1960, sección 30 y también las secciones 35, 44 y 45, que tratan, además, del problema de un enfoque lógicamente adecuado de las atribuciones de creencias). Varios de estos problemas, y otros similares concernientes al enfoque de las atribuciones de fines han sido agudamente examinados por Scheffler (1963, Parte I, sección 8).

estallido de la guerra entre Francia y Prusia en 1870. Las relaciones políticas entre las dos naciones eran tirantes por la enérgica oposición de Francia al proyecto, que por un momento pareció realizable, de colocar a un príncipe Hohenzollern en el trono de España. Bismarck había abrigado la esperanza de que este problema brindara a Prusia un *casus belli* contra Francia, pero el príncipe había renunciado a la candidatura y pareció desvanecerse la perspectiva de un conflicto militar con Francia. En estas circunstancias, un emisario francés se presentó ante el rey Guillermo de Prusia, que estaba pasando una temporada en el balneario de Ems, con la solicitud de que el rey renunciara a toda futura presentación de la candidatura. El rey se negó e informó a Bismarck del incidente con un telegrama en el que no traslucía ningún enfado, sino que simplemente le comunicaba sus razones para negarse a la solicitud. El rey dejó explícitamente a Bismarck la decisión de publicar o no el contenido del telegrama. Bismarck aprovechó la oportunidad para redactar el texto, con vistas a su publicación, de una manera calculada para inducir a Francia a declarar la guerra. Las razones que determinaron esta acción han sido discutidas por muchos autores, inclusive el mismo Bismarck.

En sus memorias,[206] Bismarck enuncia ante todo sus razones para buscar la guerra con Francia. Entre ellas figuraban su preocupación por preservar el honor nacional de Prusia; su creencia de que, en caso contrario, la pérdida de prestigio resultante obstaculizaría seriamente la formación de un imperio alemán bajo la conducción prusiana; la esperanza de que una guerra nacional contra Francia sirviera para zanjar las diferencias entre muchas de las naciones alemanas que Bismarck trataba de unir; y la información, suministrada por el jefe del Estado Mayor, de que, dado el estado de preparación militar de Prusia, no cabía esperar ninguna ventaja de postergar el estallido de la guerra. Bismarck concluye esta parte de su relato con las palabras: "Todas estas consideraciones, consciente e inconscientemente, reforzaron mi opinión de que sólo podía evitar la guerra a costa del honor de Prusia y de la fe nacional en ella. Con esta convicción, hice uso de la autorización real ... para publicar el contenido del telegrama; y ... lo reduje suprimiendo palabras, pero sin agregar ni alterar nada".[207]

La versión publicada del telegrama de Ems creó la impresión de que el rey había tratado al emisario francés de una manera insultante. En sus memorias, Bismarck expresa sinceramente las razones por las que eligió este medio para obtener sus fines: esperaba que el texto publicado "tuviera el efecto de un trapo rojo sobre el toro gálico ... Debíamos luchar ... Pero el éxito depende esencialmente de la impresión que el origen de la guerra haga sobre nosotros y los demás; es importante que nosotros seamos la parte atacada, en lo cual nos convertirá la arrogancia y la susceptibilidad gálicas si anunciamos a Europa ... que hacemos frente sin temor a las amenazas públicas de Francia".[208] La publicación del texto re-

[206] Bismarck (1899, págs. 99 y sigs.). El texto del telegrama del rey está citado en la pág. 97, y el de la versión publicada en las págs. 100-101.
[207] Bismarck (1899, pág. 100).
[208] Bismarck (1899, pág. 101).

dactado por Bismarck tuvo el efecto que éste esperaba: en París se tomó como un insulto a la nación y el gabinete francés decretó la movilización.

En cuanto a la fuerza explicativa del relato del propio Bismarck o de los que han ofrecido diversos historiadores, observamos primero que por esclarecedora que sea la expresión de razones motivantes no puede ni pretender arrojar luz sobre un aspecto muy importante de la acción de Bismarck, a saber, por qué se le ocurrió en primer término la idea de redactar el texto. En el terreno de nuestra explicación por razones, el enunciado de que se le ocurrió se ofrece simplemente como dato explicativo, como parte de la requerida especificación de los cursos de acción que el agente creía abiertos ante él. De este modo, el relato explicativo que hemos expuesto puede, a lo sumo, responder a la pregunta: dado que tal posibilidad se le ocurrió a Bismarck, ¿por qué eligió ese curso de acción?

Consideremos ahora en qué medida la explicación aquí esbozada se conforma al modelo de una acción conscientemente racional. Ante todo, presenta a Bismarck como habiendo llegado a su decisión después de un cuidadoso examen de los mejores medios disponibles para lograr su fin de provocar a Francia con el propósito de que declarara la guerra. El relato indica, además, que en la situación dada, Bismarck creía que se le abrían varios cursos de acción: publicar una versión corregida del telegrama, publicar el texto original y no publicar nada. Según su juicio, la primera alternativa, y sólo ella, tenía probabilidad de lograr el efecto deseado. Luego, si la lista de consideraciones motivantes es fácticamente correcta y completa, en el sentido de no omitir ninguna de las posibilidades contempladas por Bismarck, entonces la descripción muestra que su acción fue la de un agente conscientemente racional y que, con respecto a sus creencias y objetivos, fue racional en el sentido de uno de los criterios más simples mencionados en la sección 10.2.

Pero en realidad, no es probable que la descripción sea absolutamente completa. Por ejemplo, Bismarck debe de haber considerado, aunque fuera con brevedad, algunos cursos de acción alternativos —entre ellos, diversas maneras de corregir el texto— que no menciona en sus propias declaraciones ni en los relatos ofrecidos por otros autores que han tratado la cuestión. Los estudios disponibles sugieren que Bismarck puede haber pensado fugazmente en la posibilidad de entregar la información pertinente a todas las embajadas prusianas, pero no a la prensa para su publicación. Por ende, hay buenas razones para dudar de que las descripciones disponibles sean en realidad tan completas como sería necesario para presentar la acción de Bismarck como conscientemente racional. En defensa de las presuntas omisiones, ¿podría argüirse que dar mayor detalle sería pedante y gratuito, pues el hecho mismo de que Bismarck optara por publicar una versión corregida no basta acaso para mostrar que, aun cuando hubiera considerado otras alternativas además de las mencionadas explícitamente, las descartó por considerarlas menos promisorias? Esta es, en efecto, una manera muy plausible de defender la afirmación de que, entre todas las acciones posibles que consideró, Bismarck eligió la que según su juicio era la mejor; pero en la medida en que se acepta este argumento, se protege la racionalidad de la decisión de Bismarck incluyéndola tácitamente en

nuestra idea de sus expectativas: él no puede haber esperado mucho de las alternativas, o de lo contrario habría actuado de modo diferente.

Así, aunque en el caso del telegrama de Ems se dispone de una cantidad inusitadamente grande de información en apariencia confiable, y aunque Bismarck parece haber llegado a su decisión mediante una reflexión fría y cuidadosa, no se satisfacen por completo los rigurosos requisitos del modelo de acción conscientemente racional.

Hay otros casos que quizá se acerquen más al "ideal" del modelo. Consideremos, por ejemplo, a un ingeniero competente que busca una solución óptima para un problema de planificación cuyo ámbito de soluciones admisibles está claramente delimitado, con la probabilidad y utilidades correspondientes precisamente especificadas, y aun con una enunciación explícita del criterio de racionalidad que se debe emplear (por ejemplo, maximización de las utilidades esperadas). En este caso, puede considerarse que los objetivos y creencias que determinan la decisión del ingeniero quedan indicados de manera completa por la especificación del problema; y al aplicar al ingeniero el modelo explicativo de un agente conscientemente racional (cuyo patrón de racionalidad es el especificado en el problema dado), podemos explicar o predecir la solución o el conjunto de soluciones a que llega, que son idénticas a la óptima en teoría.

La propiedad ampliamente disposicional de la racionalidad consciente no necesita ni puede, en realidad, concebirse como una característica perdurable. Un hombre puede estar dispuesto a actuar con racionalidad consciente en algunos momentos, cuando las condiciones psicológicas y ambientales sean favorables, y sin embargo no hacerlo en otros momentos, cuando circunstancias externas perturbadoras o factores tales como la fatiga, el dolor o la preocupación por otros problemas le impiden realizar una reflexión estrictamente racional. De manera análoga, una masa de gas determinada puede comportarse "idealmente" en ciertos momentos, cuando se encuentra a elevada temperatura y baja presión, y no en otros, cuando se invierten las circunstancias.

Sin embargo, mientras que para una masa dada de gas es posible enunciar con considerable precisión las condiciones de la conducta cercana a la ideal en función de unos pocos parámetros cuantitativos, las condiciones en las cuales un individuo determinado llegará casi a actuar con racionalidad consciente sólo pueden indicarse vagamente y por medio de una lista larga y siempre abierta de categorías que incluyen tanto factores ambientales como fisiológicos y psicológicos. Hablando en términos muy amplios, el concepto de modelo explicativo de acción conscientemente racional será aplicable a aquellos casos en que el problema de decisión que el agente trata de resolver está estructurado con claridad y admite una solución relativamente simple, el agente es bastante inteligente como para hallar la solución y las circunstancias permiten una reflexión cuidadosa exenta de influencias perturbadoras.[209]

La idea de agente conscientemente racional, con su limitadísimo ámbito de aplicación, no es el único medio por el que puede darse un uso

[209] Véanse también las observaciones de Gibson (1960, págs. 165-168), que n de este punto.

explicativo y predictivo a un concepto modelo de decisión racional. Davidson, Suppes y Siegel[210] han presentado en un estudio una interesante alternativa. Estos investigadores proponen una teoría empírica de la elección inspirada en el modelo matemático de decisión con riesgo e introducen en ella la hipótesis de que las opciones de individuos humanos serán racionales en el sentido preciso de maximizar los beneficios esperados.

Como podría suponerse de antemano, se logra el riguroso carácter cuantitativo de la teoría a costa de limitar su aplicabilidad a decisiones de un tipo bastante simple, lo cual permite un control experimental estricto. En el test de la teoría ofrecida por los autores, los sujetos deben tomar una serie de decisiones, cada una de las cuales exige una elección entre dos opciones. Cada opción ofrece la perspectiva o bien de ganar una pequeña cantidad especificada de dinero, o bien de perder otra pequeña cantidad especificada, según el resultado de cierto experimento de azar, por ejemplo, echar a rodar un dado regular con marcas especiales en sus caras. Los experimentos de azar, sus posibles resultados y las ganacias y pérdidas correspondientes se describían con detalle al sujeto, quien luego realizaba su elección.

Los resultados de este experimento se conformaban bastante bien a la hipótesis de que los sujetos elegirían la opción con mayor *utilidad esperada*; se calculaba la utilidad esperada de una opción sobre la base de probabilidades y utilidades *subjetivas* postuladas teóricamente que tienen los diferentes resultados para el individuo que elige. La teoría propuesta por los autores suministra un método objetivo, aunque indirecto, para la medición simultánea e independiente de tales probabilidades y utilidades subjetivas para un agente dado. El estudio experimental demuestra que la probabilidad subjetiva que posee un resultado determinado para un sujeto dado no es, en general, igual a su probabilidad objetiva, aunque el sujeto conozca esta última; tampoco las utilidades subjetivas son proporiconales a las correspondientes ganancias o pérdidas monetarias. En realidad, normalmente una persona ignorará de manera total las probabilidades y utilidades subjetivas que tienen para él los resultados posibles, según la teoría en consideración.

Así, en la medida en que la teoría es correcta, da un sesgo muy peculiar a la idea de acción racional: aunque los sujetos realicen sus elecciones en situaciones de decisión estructuradas con claridad, actúan racionalmente (en un sentido cuantitativo definido con precisión) con respecto a las probabilidades y utilidades subjetivas que no conocen y qu, por ende, no pueden tomar en cuenta en sus reflexiones. Actúan racionalmente en el sentido de actuar *como si* trataran de maximizar las utilidades esperadas. Aquí parece haber, pues, un tipo de decisión consciente que es *inconscientemente racional* con precisión cuantitativa.

10.3.6. *La "racionalidad" de las acciones no reflexivas. La explicación por motivos inconscientes.* Muchos actos intencionales se realizan sin una anterior reflexión consciente, sin ningún cálculo de los medios que deben elegirse para el logro de un fin. Sin embargo, a menudo se explican

210 Davidson, Suppes y Siegel (1957).

tales actos en función de razones motivantes. Dray, que incluye específicamente esas explicaciones dentro del ámbito de su análisis, arguye que su concepción de la explicación racional es aplicable a cualquier acto intencional sobre la base de que "en la medida en que decimos que un acto es intencional, sea cual fuere su nivel de reflexión consciente, existe un cálculo que podría construirse para ella: el que el agente habría elaborado si hubiera tenido tiempo, si no hubiera visto de golpe qué hacer, si se le pidiera que explicara lo que hizo después de los hechos, etc. Y es haciendo explícito tal cálculo como explicamos el acto".[211]

Pero la significación explicativa de las razones o cálculos construidos de esta manera es ciertamente desconcertante. Si un agente llega a su decisión "súbita" y no por reflexión, entonces parece falso decir que la decisión puede explicarse por medio de un razonamiento que el agente podría haber hecho en circunstancias más propicias o que podría elaborar luego si se le pide que explique su acto; pues, por hipótesis, el agente no hizo ningún razonamiento similar en el momento decisivo. Las consideraciones de adecuación o racionalidad no desempeñaron ningún papel en su decisión, y una explicación en función de tales deliberaciones o cálculos sería simplemente ficticia.

Sin embargo, creo que Dray tiene cierta razón al considerar algunos actos no reflexivos como afines a los que obedecen a una reflexión cuidadosa. En efecto, las "explicaciones racionales" de tales actos pueden considerarse como explicaciones ampliamente disposicionales que aducen ciertas pautas de conducta adquiridas por el agente mediante un proceso de aprendizaje cuyas fases iniciales suponen la reflexión consciente. Consideremos, por ejemplo, el complejo conjunto de maniobras requeridas para conducir un automóvil a través de un tránsito cargado, para usar una máquina de coser o para realizar una operación quirúrgica; todas ellas se aprenden por procesos de aprendizaje que inicialmente requieren una reflexión más o menos compleja, pero que luego se realizan automáticamente, con poca o ninguna reflexión consciente, aunque a menudo de una manera que es la que el agente habría elegido si hubiera pensado en la cuestión. Por consiguiente, un acto particular de este tipo podría ser explicado, no por un cálculo elaborado que el agente de hecho no realizó, sino presentándolo como una manifestación de una disposición conductal general que el agente ha adquirido de la manera indicada.[212]

El intento de explicar una acción determinada por razones motivantes se enfrenta con otra conocida dificultad: con frecuencia desemboca en una racionalización, más que en una explicación, sobre todo cuando se basa en las razones aducidas por el agente mismo. Como observa G. Watson: "La motivación, tal como se presenta en la perspectiva de la historia, es a menudo demasiado simple y directa, y refleja la psicología de la Edad de

211 Dray (1957, pág. 123).

212 Scheffler (1963, págs. 115-116) ha sugerido, de una manera similar, que una interpretación en función de aprendizaje podría aclarar ciertos tipos de enunciados teleológicos acerca de la conducta humana. Sobre este punto, véase también el importante artículo de Suppes (1961); Gibson (1960, págs. 157-158) presenta una concepción disposicional de actos que no son deliberadamente racionales.

la Razón ... La psicología ha llegado ... a reconocer el enorme peso de los impulsos irracionales e íntimamente personales en la conducta. En la historia, la biografía y la autobiografía, sobre todo las de personajes públicos, es grande la tendencia a presentar 'buenas' razones en lugar de razones 'reales' ".[213] Por consiguiente, como continúa señalando Watson, al examinar las motivaciones de los personajes históricos es importante tomar en cuenta la significación de mecanismos psicológicos tales como la formación reactiva, "la dinámica dialéctica por la cual la tacañería se disfraza de generosidad o el pacifismo fanático que surge del intento de reprimir fuertes impulsos agresivos".[214]

La creciente comprensión de que las acciones pueden hallarse impulsadas, en considerable medida, por factores motivantes de los que el agente no tiene conciencia ha llevado a algunos historiadores a poner mucho énfasis sobre el uso más sistemático de las ideas del psicoanálisis o de teorías análogas de psicología profunda en el contexto de la explicación histórica. La alocución presidencial de W. L. Langer ante la American Historical Association [Asociación Histórica Americana] de 1957,[215] constituye una formulación y una defensa enérgicas de este programa.

Consideraciones similares han conducido a algunos filósofos que han escrito sobre motivación, a distinguir en las explicaciones de los actos de una persona entre "sus razones" para haber hecho lo que hizo y "las razones" o "las razones reales" de dichos actos.[216] En su claro estudio sobre la explicación histórica, Gardiner hace la siguiente observación a esta idea: "En general, parece correcto decir que por 'razones reales' de una persona entendemos aquellas razones que estaría dispuesta a dar en circunstancias en las que su confesión no originaría consecuencias adversas para ella. Una excepción es el uso psicoanalítico de la expresión, para el cual se adoptan criterios diferentes".[217] Pero si Gardiner tiene razón en su caracterización de lo que se entiende comúnmente por razones reales de una persona para haber actuado como lo hizo, entonces el historiador en busca de razones que expliquen correctamente las acciones humanas tendrá que renunciar a las "razones reales" en el sentido ordinario, si las investigaciones psicológicas y de otro tipo muestran que aquéllas no brindan una comprensión tan adecuada de las acciones humanas como una interpretación realizada en términos de concepciones menos familiares, incluso quizás una teoría de las motivaciones subconscientes. En efecto, Langer ha insistido vigorosamente en la necesidad de tal reorientación: "Consideradas a la luz de la moderna psicología profunda, las interpretaciones caseras y de sentido común de los historiadores de épocas pasadas,

213 Watson (1940, pág. 36).

214 *Ibíd.* Algunas sugerentes observaciones desde un punto de vista psicoanalítico sobre la noción de "racionalización" en la especificación de los motivos de una acción se encontrarán en F. Alexander (1940).

215 Langer (1958). Observaciones con un espíritu similar se hallarán en el capítulo 3 de Hughes (1964) y en la "Introducción" de Mazlish a su antología (1963) que incluye una serie de ejemplos específicos de interpretaciones inspiradas en el psicoanálisis, de temas históricos.

216 Véase, por ejemplo, Peters (1958, págs. 3-9 y *passim*).

217 Gardiner (1952, pág. 136).

aun de los más grandes, parecen lamentablemente inadecuadas, por no decir ingenuas. Evidentemente, ha llegado el momento de tomar en consideración una doctrina que tanta importancia tiene para nuestra propia disciplina".[218]

En cuanto a la noción de "razones reales" de una acción dada yo diría, pues, en primer término, que la explicación psicológica o histórica no puede depender del uso de esa noción en el lenguaje cotidiano. Pero, en segundo término, dudo de que la caracterización sugerida por Gardiner de una manera explícitamente tentativa haga justicia aun a lo que queremos significar en el lenguaje corriente cuando hablamos de las razones reales que determinaron una acción dada. En efecto, la idea de motivos subconscientes es muy común en nuestra época, por lo que estamos dispuestos a decir en el lenguaje cotidiano que las razones dadas por un agente pueden no ser las "razones reales" que provocaron su acción, aunque su declaración sea subjetivamente honesta y no tenga por qué esperar consecuencias adversas. E independientemente de que se intenten realizar explicaciones de acciones humanas en el lenguaje corriente o en los términos técnicos de alguna teoría, el criterio general para establecer lo que debe considerarse como una razón "real" y, por ende, explicativa de una acción determinada, si es que hay algo que pueda ser así considerado, no se encontrará, seguramente, examinando la manera cómo se ha usado hasta ahora la expresión "razón real", sino investigando cuál es la concepción de las razones reales que brinda la explicación más satisfactoria de la conducta humana. El uso común cambiará, en consecuencia, de manera gradual.

La estructura lógica de las explicaciones en función de motivos y procesos subconscientes es también ampliamente disposicional, en el sentido que consideramos antes: la atribución de tales motivos equivale a asignar al agente ciertas características ampliamente disposicionales, y la referencia a mecanismos subconscientes o procesos psicodinámicos refleja la admisión de leyes o principios teóricos que incluyen esas características. Pero afirmar esto no implica que todas las interpretaciones psicoanalíticas presentadas hasta ahora satisfagan los requisitos básicos de las explicaciones disposicionales científicamente adecuadas. De hecho, los criterios empíricos u operacionales para la aplicación de conceptos psicoanalíticos, y los principios teóricos en los que funcionan estos conceptos, a menudo nos son tan claros como sería deseable en interés de su aplicabilidad y testabilidad objetivas.[219] Pero no debe olvidarse que, a este respecto, las explicaciones motivacionales del sentido común a menudo tambien dejan mucho que desear y, además, que se están realizando esfuerzos por dar al psicoanálisis y a concepciones similares una forma metodológica más satisfactoria.

10.3.7. *Nota sobre los aspectos causales de las explicaciones dispo-*

[218] Langer (1958, pág. 90). Peters (1958, pág. 63) observa explícitamente que un deseo inconsciente podría constituir "la razón" de la acción de un hombre.

[219] Sobre este punto véase, por ejemplo, la crítica expuesta por Nagel (1959); y véase también la crítica y la defensa de las concepciones psicoanalíticas en otros ensayos incluidos en Hook (1959).

cionales. Se sostiene con frecuencia que las explicaciones realizadas en función de razones motivantes, habilidades aprendidas, rasgos de personalidad, etc., por tener carácter disposicional son no causales. Pero estas tesis me parece engañosa. Ante todo, como se ve en los esquemas (9.1) y (9.5), una explicación disposicional, además de la correspondiente propiedad disposicional *M*, aduce también la presencia de circunstancias —llamémosla *S*— en las que la propiedad *M* se manifestará por el síntoma —por ejemplo, una conducta de tipo *R*— cuya aparición se quiere explicar. Así, la atribución de venalidad a un agente explicará que ha cometido una traición sólo en conjunción con adecuadas suposiciones adicionales, tales como la de que se le ofreció una suma grande de dinero, lo cual, en virtud de su proclavidad venal, dio origen al acto aludido. En este caso, en analogía con el impacto de la piedra en (9.1), en el lenguaje cotidiano puede decirse que el ofrecimiento de una suma de dinero ha causado el suceso del explanandum. No puede afirmarse, pues, que las explicaciones disposicionales de este tipo no sean causales. Sin duda, la mera posesión de la propiedad disposicional *M* no se considera comúnmente como una causa; pero en tal caso la mera posesión de *M* no explica el suceso dado.

Así, cuando Gardiner observa que una explicación de la forma "*x* hizo *y* porque quería *z*" no se refiere a una relación causal entre los dos sucesos,[220] tiene razón en el sentido de que el enunciado "*x* quería *z*" no describe un suceso, sino que atribuye a *x* una propiedad ampliamente disposicional. Pero una oración "... porque ..." de la forma especificada sólo suministra una explicación en la suposición adicional de que *x* estaba en circunstancias en que, por lo menos a su entender, cabía esperar que hacer *y* condujera a *z*; y cuando se la complementa con este enunciado adicional, la explicación adopta la forma (9.5), de la cual no puede decirse que no es causal. La insistencia de Gardiner en que "las explicaciones motivacionales ... no son en modo alguno causales" [221] es provechosa en cuanto previene, como es su intención, contra la concepción de los motivos como causas fantasmales de la conducta manifiesta y contra la idea de que "en la historia nos encontramos con un mundo de 'entidades mentales' misteriosamente ocultas detrás del mundo de los cuerpos y las acciones físicas, separadas de él y que sin embargo lo controlan; [222] pero corre el riesgo de oscurecer las estrechas semejanzas que hemos señalado entre las explicaciones motivacionales y otras explicaciones consideradas, por lo general, como causales.[223]

220 Gardiner (1952, pág. 124).

221 Gardiner (1952, págs. 133-134). Véase también la opinión de Ryle de que "explicar una acción como debida a un motivo o inclinación especificados no es describir la acción como efecto de una causa especificada. Los motivos no son sucesos y, por ende, no son adecuados como causas" (1949, pág. 113).

222 Gardinier (1952, pág. 51).

223 En este contexto, véase también el sugerente examen de disposiciones, razones y causas en Dray (1957, págs. 150-155). En contraste con la opinión de que "sólo los sucesos y procesos pueden ser causas" (pág. 151), Dray sostiene que una "característica disposicional es un tipo de 'condición permanente'; y las condiciones permanentes pueden ser causas tanto como las precipitantes" (pág. 152). La tesis

11. *Observaciones finales*

Al comienzo de este capítulo contrapusimos las preguntas sobre el porqué que piden razones y las preguntas sobre el porqué que piden explicaciones. Las primeras solicitan fundamentos que hagan *creíbles* ciertos enunciados empíricos; las segundas solicitan una información que explique ciertos hechos y los haga, así, *inteligibles*. Nos hemos dedicado a examinar las maneras como la ciencia responde a las preguntas sobre el porqué del último tipo y a caracterizar el género de comprensión que ella brinda.

Observamos que la explicación científica no pretende crear una sensación de familiaridad con el explanandum; la "reducción a lo familiar" es a lo sumo un aspecto incidental de ella. La comprensión que transmite reside más bien en la aprehensión de que el explanandum se adecua a un sistema de uniformidades representado por leyes empíricas o principios teóricos, o puede subsumirse en éstos. Según el carácter lógico de las uniformidades, tal subsunción será deductiva o inductiva, en un sentido que aspiran a hacer explícito nuestros dos modelos básicos.

Quiero destacar aquí una vez más que hay profundas diferencias lógicas entre esos dos modos de explicación. No se trata de que en una explicación estadística la oración explanandum presente una cláusula modal tal como "probablemente" o "casi con certeza"; el explanandum es una oración no modal tanto en la explicación y la predicción probabilística como en las deductivo-nomológicas. Pero en la explicación inductivo-estadística, a diferencia de su análoga deductiva, el explanans sólo hace al explanandum más o menos probable y no lo implica con certeza deductiva. Otra diferencia, que hasta ahora no parece haber recibido atención, reside en lo que llamo la relatividad epistémica de la explicación probabilística, es decir, el hecho de que podamos hablar con sentido de una explicación probabilística, aun potencial, sólo con respecto a una clase *K* de enunciados que repesentan una particular situación cognoscitiva. El concepto de explicación deductivo-nomológica no exige tal relativización.

Ejemplificamos e hicimos explícito el papel explicativo de presuntas leyes y principios teóricos mediante un análisis de diversos tipos de explicaciones ofrecidas en diferentes campos de la ciencia empírica. Este panorama no pretende ser completo; se lo podría haber ampliado con el

de que la explicación por razones es "una especie de explicación causal común" ha sido defendida de manera muy interesante, sobre fundamentos diferentes de los presentados aquí, por Davidson (1963), donde también se examinan una cantidad de objeciones. Debe recordarse que la concepción común de la explicación causal es bastante estrecha y vaga, y que —al menos en la física— ha sido reemplazado por la concepción más general y precisa de explicación por medio de una teoría determinista. Lo ilustra el caso, considerado en la sección 2, de la teoría newtoniana del movimiento y de la gravedad: dado el "estado" de un sistema cerrado de masas puntuales en un momento definido, la teoría determina el estado del sistema en cualquier otro momento y, por lo tanto, permite explicar un estado particular del sistema con referencia a un estado anterior. Los términos de la relación causal consisten aquí no en sucesos, sino en *estados* momentáneos del sistema, representados por las masas, posiciones y velocidades de las partículas constituyentes en el momento en cuestión.

examen del uso explicativo de conceptos y teorías tipológicos, del análisis funcional, de ideas psicoanalíticas, etcétera.[224]

La tesis central de este ensayo fue, dicho brevemente, que toda explicación científica supone, de manera explícita o por implicación, una subsunción de su ámbito bajo regularidades generales; que trata de suministrar una comprensión sistemática de los fenómenos empíricos mostrando que éstos entran en un nexo nómico. Esta concepción, que ha sido expuesta en detalle en las secciones anteriores, no pretende simplemente describir las explicaciones que se brindan de hecho en las ciencias empíricas, pues, para no mencionar más que una razón, no existe un acuerdo bastante claro y generalmente aceptado acerca de lo que constituye una explicación científica. La concepción aquí expuesta tiene, más bien, la naturaleza de una *explicación*, que pretende reemplazar una noción familiar pero vaga y ambigua por otra caracterizada con mayor precisión y que sea sistemáticamente fructífera y esclarecedora. En realidad, nuestro análisis explicativo ni siquiera nos ha conducido a una cabal definición de un concepto-"explicatum" preciso de la explicación científica; sólo tendió a hacer explícitos algunos aspectos especialmente importantes de tal concepto.[225]

Como cualquier otra explicación, la concepción aquí expuesta debe justificarse mediante argumentos adecuados. En nuestro caso, éstos deben mostrar que el enfoque propuesto hace justicia a las explicaciones que se conviene, en general, en considerar como casos de explicación científica, y que suministra una base para llevar a cabo un análisis lógico y metodológico sistemáticamente fructífero de los procedimientos explicativos utilizados en las ciencias empíricas. Esperamos que los argumentos presentados en este capítulo hayan alcanzado dicho objetivo.

BIBLIOGRAFIA

Alexander, F.: "Psychology and the Interpretation of Historical Events". En Ware (1940), 48-57.

Alexander, H. G.: "General Statements as Rules of Inference?" En Feigl, Scriven, y Maxwell (1958), 309-29.

Arrow, K. J.: "Mathematical Models in the Social Sciences". En Lerner, D. y H. D. Laswell(comps.): *The Policy Sciences.* Stanford, Stanford University Press, 1951, 129-54.

224 La lógica del análisis funcional es estudiada en el capítulo X; los conceptos y teorías tipológicas, en la Tercera Parte, *Estructura y función de los conceptos científicos.* Kahl (1963) ofrece una interesante y útil colección de explicaciones tomadas de la física, la biología, la psicología y la historia.

225 En la sección 5 del capítulo X, "La lógica de la explicación" y en su Postscriptum se aclara que una caracterización más completa de este concepto y, a fortiori, una definición explicativa completa plantea problemas adicionales. Otro problema que se plantea es el mencionado en la nota 33 de ese mismo capítulo, tal como ha sido reimpreso en este volumen.

Baernstein, H. D. y Hull, C. L.: "A Mechanical Model of the Conditioned Reflex". *The Journal of General Psychology* 5 (1931), 99-106.
Barker, S. F.: *Induction and Hypothesis.* Ithaca, N.Y., Cornell University Press, 1957.
Barker, S. F.: "The Role of Simplicity in Explanation". En Feigl y Maxwell (1961), 265-74.
Bartley, W. W.: "Achilles, the Tortoise, and Explanation in Science and History". *The British Journal for the Philosophy of Science* 13 (1962), 15-33.
Baumol, William J.: *Economic Theory and Operations Analysis.* Englewood Cliffs, N.J., Prentice-Hall, 1961.
Baumrin, B. (comp.): *Philosophy of Science. The Delaware Seminar.* Volume I, 1961-62. Nueva York, John Wiley & Sons, 1963.
Beale, H. K.: "What Historians Have said About the Causes of the Civil War". En *Theory and Practice in Historical Study: A Report of the Committee on Historiography,* Social Science Research Council, Bulletin 54; Nueva York (1964), 53-92.
Bertalanffy, L. von: *Modern Theories of Development.* Londres, Oxford University Press, 1933.
Bertalanffy, L. von: "Problems of General System Theory". *Human Biology* 23 (1951), 302-12.
Bertalanffy, L. von: "General System Theory". En Bertalanffy, L. von, y A. Rapoport, (comps.) *General Systems. Yearbook of the Society for the Advancement of General Systems Theory.* Vol. I, 1956.
Bismarck, Otto von: *Bismarck. The Man and the Statesman: Being the Reflections and Reminiscences of Otto, Prince von Bismarck.* Traducido del alemán bajo la supervisión de A. J. Butler. Vol. II. Nueva York, Harper and Row, 1899.
Bochmer, H.: *Luther and the Reformation in the Light of Modern Research.* Traducido por E. S. G. Potter. Nueva York, The Dial Press, 1930.
Boltzmann, L.: *Vorlesungen über Maxwells Theorie der Elektrizität und des Lichtes.* I. Theil. Leipzig, Barth, 1891.
Boltzmann, L.: *Populäre Schriften.* Leipzig, Barth, 1905.
Bonhoeffer, K. F.: "Über physikalisch-chemische Modelle von Lebensvorgängen". *Studium Generale* 1 (1948), 137-43.
Bondi, H.: *The Universe at Large.* Londres, Heinemann, 1961.
Braithwaite, R. B.: *Scientific Explanation.* Cambridge, England, Cambridge University Press, 1953.
Brandt, R. y J. Kim: "Wants as Explanations of Actions". *The Journal of Philosophy* 60 (1963), 425-35.
Bridgman, P. W.: *The Logic of Modern Physics.* Nueva York, Macmillan, 1927.
Brodbeck, May: "Models, Meaning, and Theories". En Gross (1959), 373-403.
Brodbeck, May: 'Explanations, Predictions, and 'Imperfect' Knowledge". En Feigl y Maxwell (1962), 231-72
Bromberger, S.: "The Concept of Explanation". Tesis de doctorado, Harvard University, 1960.
Bromberger, S.: "An Approach to Explanation". En Butler, R. (comp.), *Studies in Analytical Philosophy.* Oxford, Blackwell (en prensa).
Bush, R. R. y F. Mosteller: *Stochastic Models for Learning.* Nueva York, John Wiley & Sons, 1955.
Campbell, N. R.: *Physics: The Elements.* Cambridge, England, Cambridge University Press, 1920.
Campbell, N. R.: *What is Science?* Nueva York, Dover, 1952. (1a. ed. 1921.)
Carington, W.: *Matter, Mind and Meaning.* Londres, Methun, 1949.
Carnap, R.: "Testability and Meaning". *Philosophy of Science* 3 (1936) y 4 (1937). Reproducido parcialmente en Feigl y Brodbeck (1953).
Carnap, R.: *The Logical Syntax of Language.* Nueva York, Harcourt, Brace and World, 1937.
Carnap, R.: "Logical Foundations of the Unity of Science". En *International Encyclopedia of Unified Science,* Vol. I, Nº 1. Chicago, University of Chicago Press, 1938. Reproducido en Feigl y Sellars (1949), 408-23.
Carnap, R.: "On inductive Logic". *Philosophy of Science* 12 (1945), 72-97.

Carnap, R.: *Logical Foundations of Probability*. Chicago, University of Chicago Press, 1950; 2ª ed. revisada 1962. Citada en este cap. como Carnap (1950).

Carnap, R.: "Inductive Logic and Science". *Proceedings of the American Academy of Arts and Sciences,* vol. 80 (1951-54), 187-97.

Carnap, R.: *The Continuum of Inductive Methods.* Chicago, University of Chicago Press, 1952.

Carnap, R.: "The Methodological Character of Theoretical Terms". En Feigl y Scriven (1956), 38-76.

Carnap, R.: "The Aim of Inductive Logic". En Nagel, Suppes y Tarski (1962) 303-18.

Chisholm, R.: "Sentences about Believing". En Feigl, Scriven y Maxwell (1958) 510-20.

Churchman, C. W.: *Prediction and Optimal Decision.* Englewood Cliffs, N. J., Prentice-Hall, 1961.

Cohen, M. R. y E. Nagel: *An Introduction to Logic and Scientific Method.* Nueva York, Harcourt, Brace & World, 1934.

Conant, James B.: *Science and Common Sense.* New Haven, Yale University Press, 1951.

Craig, W. "Replacement of Auxiliary Expressions". *Philosophical Review* 65 (1956) 38-55.

Cramér, H.: *Mathematical Methods of Statistics.* Princeton, Princeton University Press, 1946.

Danto, A. C.: "On Explanations in History". *Philosophy of Science* 23 (1956) 15-30.

Davidson, D.: "Actions, Reasons, and Causes". *The Journal of Philosophy* 60 (1963) 685-700.

Davidson, D., P. Suppes y S. Siegel: *Decision Making: An Experimental Approach.* Stanford, Stanford University Press, 1957.

Dewey, John: *How We Think.* Boston, D. C. Heath & Co., 1910.

Donagan, A.: "Explanation in History". *Mind* 66 (1957) 145-64. Reproducido en Gardiner (1959), 428-43.

Dray, W.: "Explanatory Narrative History". *Philosophical Quarterly* 4 (1954) 15-27.

Dray, W.: *Laws and Explanation in History.* Oxford, Oxford University Press, 1957.

Dray, W.: "'Explaining What' in History". En Gardiner (1959) 403-08.

Dray, W.: "The Historical Explanation of Actions Reconsidered". En Hook (1963) 105-35.

Duhem, P.: *La Théorie Physique. Son Objet et Sa Structure.* París, Chevalier et Rivière, 1906. (También traducido por P. P. Wiener con el título *The Aim and Structure of Physical Theory.* Princeton, Princeton University Press, 1954).

Feigl, H.: "Some Remarks on the Meaning of Scientific Explanation". En Feigl y Sellars (1949) 510-14.

Feigl, H.: "Notes on Causality". En Feigl y Brodbeck (1953) 408-18.

Feigl. H. y M. Brodbeck (comps.): *Readings in the Philopsophy of Science.* Nueva York, Appleton-Century-Crofts, 1953.

Feigl, H. y G. Maxwell (comps.): *Current Issues in the Philosophy of Science.* Nueva York, Holt, Rinehart & Winston, 1961.

Feigl, H. y M. Scriven (comps.): *Minnesota Studies in the Philosophy of Science,* Vol. III. Mineapolis, University of Minnesota Press, 1962.

Feigl, H. y M. Scriven (comps.): *Minnesota Studies in the Philopsophy of Science,* Vol. I. Minneapolis, University of Minnesota Press, 1956.

Feigl, H., M. Scriven y G. Maxwell (comps.): *Minnesota Studies in the Philosophy of Science,* Vol. II. Mineapolis, University of Minnesota Press, 1958.

Feigl, H. y W. Sellars (comps.): *Readings in Philosophical Analysis.* Nueva York, Appleton-Centry-Crofts, 1949.

Feyerabend, P. K.: "Explanation, Reduction, and Empiricism". En Feigl y Maxwell (1952) 28-97.

Feyerabend, P. K.: Resumen de Hanson (1963) en *Philosophical Review* 73 (1964) 264-66.

Frank, P.: *Philosophy of Science.* Englewood Cliffs, N.J., Prentice-Hall, 1957.

Frankel, C.: "Explanation and Interpretation in History". En Gardiner (1959) 408-27. Reproducido de *Philosophy of Science* 24 (1957) 137-55.

French, T. M.: *The Integration of Behavior.* Vol. I. *Basic Postulates.* Chicago, University of Chicago Press, 1952.
Freud, S.: *Psychopathology of Everyday Life.* Traducido por A. A. Brill. Nueva York, The New American Library (Mentor Book Series), 1951.
Galilei, Galileo: *Dialogues Concerning Two New Sciences.* Traducido por H. Crew y A. de Salvio. Evanston, Northwestern University, 1946.
Gallie, W. B.: "Explanation in History and the Genetic Sciences". *Mind* 64 (1955). Reproducido en Gardiner (1959) 368-402.
Gardiner, P.: *The Nature of Historical Explanation.* Oxford, Oxford University Press, 1952.
Gardiner, P. (comp.): *Theories of History.* Nueva York, The Free Press, 1959.
Gasking, D.: "Causation and Recipes". *Mind* 64 (1955) 479-87.
Gauss, C. F.: "Allgemeine Lehrsaetze in Beziehung of die im verkehrten Verhaeltnisse des Quadrats der Entfernung wirkenden Anziehungs- and Abstossungs-Kraefte". (Publicado 1840.) Reproducido en *Ostwalds Klassiker der exacten Wissenschaften,* Nº 2, Leipzig, Wilhelm Engelmann, 1889.
Gibson, Q.: *The Logic of Social Enquiry.* Londres, Routledge and Kegan Paul; Nueva York: Humanities Press, 1960.
Goldstein, L. J.: "A Note on the Status of Historical Reconstructions". *The Journal of Philosophy* 55 (1958) 473-79.
Goodman, Nelson: "The Problem of Counterfactual Conditionals". *The Journal of Philosophy* 44 (1947) 113-28. Reproducido con pequeños cambios, como primer capítulo de Goodman (1955).
Goodman, Nelson: *Fact, Fiction, and Forecast.* Cambridge, Mass., Harvard University Press, 1955.
Goudge, T. A.: "Causal Explanation in Natural History". *The British Journal for the Philosophy of Science* 9 (1958) 194-202.
Gross, L. (comp.): *Symposium on Sociological Theory.* Nueva York, Harper & Row, 1959.
Grünbaum, A.: "Temporally Asymmetric Principles, Parity between Explanation and Prediction, and Mechanism vs. Teleology". En Baumrin (1953) 57-96.
Grünbaum, A.: *Philosophical Problems of Space and Time.* Nueva York, Knopf, 1963a.
Hanson, N.: "On the Symmetry between Explanation and Prediction". *The Philosophical Review* 68 (1959) 349-58.
Hanson, N. R.: *The Concept of the Positron. A Philosophical Aanalysis.* Cambridge, England, Cambridge University Press, 1963.
Helmer, O. y P. Oppenheim: "A Syntactical Definition of Probability and of Degree of Confirmation". *The Journal of Symbolic Logic* 10 (1945) 25-60.
Helmer, O. y N. Rescher: "On the Epistemology of the Inexact Sciences". *Management Science* 6 (1959).
Hempel, C. G.: "The Function of General Laws in History". *The Journal of Philosophy* 39 (1942) 35-48. Reproducido en este volumen.
Hempel, C. G.: "Studies in the Logic of Confirmation". *Mind* 54 (1945) 1-26 y 97-121. Reproducido en este volumen.
Hempel, C. G.: "A Note on the Paradoxes of Confirmation". *Mind* 55 (1946) 79-82.
Hempel, C. G.: "Problems and Changes in the Empiricist Criterion of Meaning". *Revue Internationale de Philosophie,* No. 11 (1950) 41-63.
Hempel, C. G.: "The Concept of Cognitive Significance: A Reconsideracion." *Proceedings of the American Academy of Arts and Sciences,* Vol. 80, No. 1 (1951) 61-77.
Hempel, C. G.: "General System Theory and the Unity of Science". *Human Biology* 23 (1951a) 313-27.
Hempel, C. G.: "The Theoretician's Dilemma". En Feigl, Scriven y Maxwell (1958) 37-98. Reproducido en este volumen.
Hempel, C. G.: "Empirical Statements and Falsifiability". *Philosophy* 33 (1958a) 342-48.
Hempel, C. G.: "The Logic of Functional Analysis". En Gross (1959) 271-307. Reproducido en este volumen.
Hempel, C. G.: "Inductive Inconsistencies". *Synthese* 12 (1960) 439-69. Reproducido en este volumen.

Hempel, C. G.: "Deductive-Nomologival *vs.* Statistical Explanation". En Feigl y Maxwell (1962) 98-169.
Hempel, C. G. y P. Oppenheim: "A Definition of 'Degree of Confirmation'". *Philosophy of Science* 12 (1945) 98-115.
Hempel, C. G. y P. Oppenheimer: "Studies in the Logic of Explanation". *Philosophy of Science* 15 (1948) 135-75. Reproducido en este volumen.
Henson, R. B.: "Mr. Hanson on the Symmetry of Explanation and Prediction". *Philosophy of Science* 30 (1963) 60-61.
Hertz, H.: *Die Prinzipien der Mechanik.* Leipzig, Johann Ambrosius Barth, 1894.
Hesse, Mary B.: *Models and Analogies in Science.* Londres y Nueva York, Sheed and Ward, 1963.
Homans, George C.: *Social Behavior. Its Elementary Forms.* Nueva York, Harcourt, Brace & World, 1961.
Hook, S. (comp.): *Psychoanalysis, Scientific Method, and Philosophy.* Nueva York, New York University Press, 1959.
Hook, S. (comp.): *Philosophy and History.* Nueva York, New York University Press, 1963.
Hughes, H. S.: *History as Art and Science.* Nueva York, Harper & Row, 1964.
International Union of History and Philosophy of Sciences: *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences.* Dordrecht, Holland, D. Reidel, 1961.
Kahl, R. (comp.): *Studies in Explanation. A Reader in the Philosophy of Science.* Englewood Cliffs, N. J., Prentice-Hall, 1963.
Kemeny, J. G., y P. Oppenheim: "Degree of actual Support". *Philosophy of Science*, 19 (1952) 307-24.
Kemeny, J. G., y P. Oppenheim: "On Reduction". *Philosophical Studies*, 7 (1956) 6-19.
Keynes, J. M.: *A Treatise on Probability.* Londres, Macmilan, 1921.
Kim, J.: "Explanation, Prediction, and Retrodiction: Some Logical and Pragmatic Considerations". Tesis de doctorado, Princeton University, 1962.
Körner, S. (comp.): *Observation and Interpretation: Proceedings of the Ninth Symposium of the Colston Research Society.* Nueva York, Academic Press, y Londres, Butterworths Scientific Publications, 1957.
Kolmogooff, A.: *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlín, Springer, 1933.
Krueger, R. G., y Hull, C. L.: "An Electro-Chemical Parallel to the Conditioned Reflex". *The Journal of General Psychology*, 5 (1931) 262-69.
Langer, W. L.: "The Next Assignment". *The American Historical Review*, 63 (1958) 283-304. Reproducido en Mazlish (1963) 87-107.
Lazarsfeld, P. F.: "The American Soldier—An Expository Review". *Public Opinion Quarterly*, 13 (1949) 377-404.
Lazarsfeld, P. F.: *Mathematical Thinking in the Social Sciences.* Nueva York, The Free Press, 1954.
Leduc, S.: *The Mechanism of Life.* Traducido por W. D. Butcher, Nueva York, Rebman Co., 1911.
Leduc, S.: *La Biologie Synthétique.* París, 1912.
Lewis, C. I.: *An Analysis of Knowledge and Valuation.* La Salle, Ill., Open Court Publishing Co., 1946.
Lodge, Sir O.: *Modern Views of Electricity.* Londres, Macmillan, 1889.
Luce, R. D., y H. Raiffa: *Games and Decisions.* Nueva York, John Wiley, 1957.
Mandelbaum, M.: "Historical Explanation: The Problem of 'Covering Laws'". *History and Theory*, 1 (1961) 229-42.
Mandler, G., y W. Kessen: *The Language of Psychology.* Nueva York, John Wiley & Sons, 1959.
Margenau, H.: *The Nature of Physical Reality.* Nueva York, McGraw-Hill, 1950.
Mazlish, B. (comp.): *Psychoanalysis and History.* Englewod Cliffs, N. J., Prentice-Hall, 1963.
Maxwell, J. C.: "On Faraday's Lines of Force". *Transactions of the Cambridge Philosophical Society*, 10 (1864) 27-83.
Mendel, A.: "Evidence and Explanation". En *Report of the Eighth Congress of the International Musicological Society, New York, 1961.* La Rue, Jan (comp.). Kassel, Bärenreiter-Verlag, 1962. Vol. II, 3-18.

Miller, N. E.: "Comments on Theoretical Models. Illustrated by the Development of a Theory of Conflict Behavior". *Journal of Personality*, 20 (1951) 82-190.
Mises, R. von: *Wahrscheinlichkeitsrechnung und ihre Anwendungen in der Statistik und theoretischen Physik*. Viena, 1931. Reproducido Nueva York, M. S. Rosenberg 1945.
Mises, R. von: *Probability, Statistics and Truth*. Londres, William Hodge & Co., 1939.
Mises, R. von: *Positivism. A Study in Human Understanding*. Cambridge, Mass., Harvard University Press, 1951.
Moulton, F. R., y J. R. Schifferes: *The Autobiography of Science*. Garden City, N.Y., Doubleday & Co., 1945.
Muir, R.: *A Short History of the British Commonwealth*. Vol. II. Londres, George Philip and Son, 1922.
Nagel, E.: *Principles of the Theory of Probability*. Chicago, University of Chicago Press, 1939.
Nagel, E.: *Logic without Metaphysics*. Nueva York, The Free Press, 1956.
Nagel, E.: "Methodological Issues in Psychoanalytic Theory". En Hook (1959) 38-56.
Nagel, E.: *The Structure of Science: Problems in the Logic of Scientific Explanation*. Nueva York, Harcourt, Brace & World, Inc., 1961.
Nagel, E., P. Suppes y A. Tarski (comp.): *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*. Stanford, Stanford University Press, 1962.
Neumann, J. von y O. Morgenstern: *Theory of Games and Economic Behavior*. Princeton, Princeton University Press, 2ª ed., 1947.
Passmore, J.: "Law and Explanation in History". *The Australian Journal of Politics and History*, 4 (1958) 269-76.
Passmore, J.: "Explanation in Everyday Life, in Science, and in History". *History and Theory*, 2 (1962) 105-23.
Peters, R. S.: *The Concept of Motivation*. Londres, Routledge and Kegan Paul; Nueva York, Humanities Press, 1958.
Pitt, J.: "Generalizations in Historical Explanation". *The Journal of Philosophy*, 56 (1959) 578-86.
Popper, K. R.: *Logik der Forschung*. Viena, Springer, 1935.
Popper, K. R.: "The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory". En Körner (1957) 65-70.
Popper, K. R.: "The Aim of Science". *Ratio*, 1 (1957a) 24-35.
Popper, K. R.: *The Logic of Scientific Discovery*. Londres, Hutchinson, 1959.
Popper, K. R.: *Conjectures and Refutations*. Nueva York, Basic Books (1962).
Price, H. H.: "The Theory of Telepathy". *Horizon*, 12 (1945) 45-63.
Quine, W. V. O.: *Word and Object*. Coedición de Technology Press of the Massachusetts Institute of Technology y John Wiley and Sons, Nueva York, 1960.
Ramsey, F. P.: *The Foundations of Mathematics and Other Logical Essays*. Londres, Routledge and Kegan Paul; Nueva York, Harcourt, Brace & World, 1931.
Reichenbach, H.: *The Theory of Probability*. Berkeley y Los Angeles, The University of California Press, 1949.
Rescher, N.: "A Theory of Evidence". *Philosophy of Science*, 25 (1958) 83-94.
Rescher, N.: "Discrete State Systems, Markov Chains, and Problems in the Theory of Scientific Explanation and Prediction". *Philosophy of Science*, 30 (1963) 325-45.
Russell, S. B.: "A Practical Device to Simulate the Working of Nervous Discharges". *The Journal of Animal Behavior*, 3 (1913) 15-35.
Ryle, G.: *The Concept of Mind*. Londres, Hutchinson, 1949.
Ryle, G.: "'If, So', and 'Because'". En Black, M. (comp.), *Philosophical Analysis*. Ithaca, N.Y., Cornell University Press, 1950.
Savage, L. J.: *The Foundations of Statistics*. Nueva York, John Wiley & Sons, 1954.
Scheffler, I.: "Explanation, Prediction, and Abstraction". *The British Journal for the Philosophy of Science*, 7 (1957) 293-309.
Scheffler, I.: *The Anatomy of Inquiry: Philosophical Studies in the Theory of Science*. Nueva York, Alfred A. Knopf, 1963.
Schlick, M.: "Die Kausalität in der gegenwärtigen Physik". *Die Naturwissenschaften* 19 (1931). Traducido por D. Rynin, "Causality in Contemporary Physics". *The British Journal for the Philosophy of Science*, 12 (1962) 177-93 y 281-98.

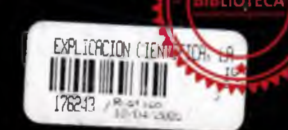
Schwiebert, E. G.: *Luther and His Times.* St. Louis, Concordia Publishing House, 1950.
Sciama, D. W.: *The Unity of the Universe.* Garden City, N. Y., Doubleday and Co. (Anchor Books), 1961.
Scriven, M.: "Definitions, Explanations, and Theories". En Feigl, Scriven y Maxwell (1958) 99-195.
Scriven, M.: "Truisms as the Grounds for Historical Explanations". En Gardiner (1959) 443-75.
Scriven, M.: "Explanation and Prediction in Evolutionary Theory". *Science,* 130 (1959a) 477-82.
Scriven, M.: "Explanations, Predictions, and Laws". En Feigl and Maxwell (1962) 170-230.
Scriven, M.: "The Temporal Asymmetry between Explanations and Predictions". En Baumrin (1963) 97-105.
Scriven, M.: "New Issues in the Logic of Explanation". En Hook (1963a) 339-61.
Seeliger, R.: "Analogien und Modelle in der Physick". *Studium Generale,* 1 (1948) 125-37.
Sellars, W.: "Inference and Meaning". *Mind,* 62 (1953) 313-38.
Sellars, W.: "Counterfactuals, Dispositions, and the Causal Modalities". En Feigl, Scriven y Maxwell (1958) 225-308.
Society for Experimental Biology: *Models and Analogues in Biology: Symposia of the Society for Experimental Biology, Number XIV.* Cambridge, England, Cambridge University Press, 1960.
Suppes, P.: "The Philosophical Relevance of Decision Theory". *The Journal of Philosophy,* 58 (1961) 605-14.
Svedberg, T.: *Die Existenz der Moleküle.* Leipzig, Akademische Verlegsgesellschaft, 1912.
Thomson, Sir William: *Notes of Lectures on Molecular Dynamics and the Wave Theory of Light.* Baltimore, The Johns Hopkins University, 1884.
Tolman, E. C.: "A Psychological Model". En Parsons, T. y E. A. Shils (comps.): *Toward a General Theory of Action.* Cambridge, Mass., Harvard University Press, 1951, 277-361.
Toulmin, S.: *The Philosophy of Science.* Londres, Hutchinson, 1953.
Toulmin, S.: *The Uses of Argument.* Cambridge, England, Cambridge University Press, 1958.
Toulmin, S.: *Foresight and Understanding.* Londres, Hutchinson, 1961; Nueva York, Harper & Row (Torchbook), 1963.
Toynbee, A.: *The World and the West.* Londres, Oxford University Press, 1953.
Turner, J.: "Maxwell on the Method of Physical Analogy". *The British Journal for the Philosophy of Science,* 6 (1955) 226-38.
Turner, J.: "Maxwell on the Logic of Dynamical Explanation". *Philosophy of Science* 23 (1956) 36-47.
Ware, C. F. (comp.): *The Cultural Approach to History.* Nueva York, Columbia University Press, 1940.
Watkins, W. H.: *On Understanding Physics.* Cambridge, England, Cambridge University Press, 1938.
Watson, G.: "Clio and Psyche: Some Interrelations of Psychology and History". En Ware (1940) 34-47.
Weber, Max: *On the Methodology of the Social Sciences.* Traducido y compilado por Shils, E. A. y H. A. Finch. Nueva York, The Free Press, 1949.
Weingartner, R. H.: "The Quarrel about Historical Explanation". *The Journal of Philosophy,* 58 (1961) 29-45.
Wiener, N.: *Cybernetics.* Nueva York, John Wiley & Sons, 1948.
Williams, D. C.: *The Ground of Induction.* Cambridge, Mass., Harvard University Press, 1947.

La explicación científica
Carl G. Hempel

Paidós
Básica

Si bien las ciencias evidencian un desarrollo casi cotidiano e innegable, adolecen de una limitación que, sin embargo, no altera la efectividad de nuestros descubrimientos: no pueden dar cuenta de la validez formal de sus procedimientos. Es decir, pueden explicar, mas no pueden explicarse, y ésta es una explicación que ha quedado a cargo de la filosofía de la ciencia. Pues bien, es en este ámbito donde se desarrolla el presente libro, que ofrece a su vez cuatro grandes núcleos temáticos: coafirmación, inducción y creencia racional; concepciones de la significación cognitiva; estructura y función de los conceptos y teorías científicos; y explicación científica.
El autor, de este modo, explora los aspectos pragmático-psicológicos y sistemático-lógicos de la explicación en las ciencias naturales y sociales y en la historiografía, presentando dos modelos básicos de hechos explicativos —el deductivo y el inductivo-probabilístico—, analizando su validez y su fin y examinando en qué medida iluminan la estructura y el fundamento de los distintos tipos de explicación que se utilizan en los diversos campos de la investigación. El resultado es un texto que interesará tanto a estudiantes y profesionales de la filosofía como a sociólogos, naturalistas e historiadores preocupados por los aspectos metodológicos y filosóficos de sus disciplinas.
Carl G. Hempel, investigador alemán radicado en los Estados Unidos y conocido como uno de los principales exponentes del positivismo lógico, es también autor de libros como *Filosofía de la ciencia natural* y *Fundamentos de la formación de conceptos en ciencia empírica.*





Diseño: Mario Eskenazi